VOLUME VIII, No. 1 /1963-64 November, 1965 -66

①

EDUCATIONAL RESEARCH

A Review for Teachers and all concerned with Progress in Education

Published for
THE NATIONAL FOUNDATION FOR EDUCATIONAL
RESEARCH IN ENGLAND AND WALES

by
NEWNES EDUCATIONAL PUBLISHING CO. LTD.
Tower House, Southampton Street, London, W.C.2

EDUCATIONAL RESEARCH

Editor:
W. D. WALL
assisted by
D. A. PIDGEON
A. E. SANDERSON
K. M. MILLER

Educational Research is published for the National Foundation by Newnes Educational Publishing Co., Ltd., once a term, in November, February and June. The Annual subscription as from November 1965 is 23/-, individual issues 7/8d. (including postage, U.K. only).

Subscriptions, orders and advertisements should be sent to Newnes Educational Publishing Co., Ltd., Tower House, Southampton Street, London, W.C.2.

Communications on general editorial matters including MSS. for publication, books and other printed matter for review or abstracting should be sent to the Director, N.F.E.R., The Mere, Upton Park, Slough, Bucks.

Note to contributors: Contributions which make a readable synthesis of valid research on the practical problems posed by the English educational system will be particularly welcome. The length of such contributions should be from 4,000 to 7,000 words, though shorter notes will be considered. Abstracts of theses or similar material should be of sufficient length and detail to acquaint the reader with the main outcomes of the work, the sampling on which it has been based and the experimental, statistical and other techniques used. Such synopses should be submitted in agreement with the head of the university department concerned.

All papers submitted for publication should conform to the conventional usage of the review and especially to the form in which bibliographical references are given. They should be typed (double space) on one side of quarto paper.

Intending contributors would be well advised in the first instance to send to the editor a brief synopsis of their paper with the elements of a bibliography.

In certain cases it may be possible for the Foundation's Information Service to aid in the gathering of relevant research material.

Only material which has not been published already elsewhere can normally be accepted for publication. The copyright in any material published is retained by the Editor. Reproduction in whole or in part is forbidden except (i) under the terms of the Fair Copying Declaration of the Royal Society or (ii) when specific permission has been given by the Editor.

Contributors receive fifteen off-prints of their article and five free copies of the number in which it appears. Further off-prints may be ordered at cost price when the proofs are returned to the Editor.

The conclusions drawn and the opinions expressed in papers are those of the authors. They should not be taken to represent the view of the N.F.E.R.

Volume VIII, No. 1

November, 1965

EDUCATIONAL RESEARCH

A Review for
Teachers and all concerned with
Progress in Education

Published for
THE NATIONAL FOUNDATION FOR EDUCATIONAL
RESEARCH IN ENGLAND AND WALES

by

NEWNES EDUCATIONAL PUBLISHING CO. LTD.
Tower House, Southampton Street, London, W.C.2

SUBSCRIPTION RATE

In the issue of *Educational Research* for June 1965, it was announced that a new Subscription rate would come into operation with the publication of the present issue.

The Annual Subscription rate for the journal is now 23/-, including postage in the United Kingdom. The price for individual issues has been fixed at 7/8d, including postage.

© *National Foundation for Educational Research in England and Wales*

Printed by King, Thorne & Stace Limited, School Road, Hove, Sussex
and published by Newnes Educational Publishing Co., Ltd., Tower House, Southampton Street, London, W.C.2.,
for the National Foundation for Educational Research in England and Wales, The Mere, Upton Park, Slough, Bucks.

DATE OF BIRTH AND SCHOLASTIC PERFORMANCE

by DOUGLAS A. PIDGEON
Deputy Director, National Foundation for Educational Research

THE fact that intellectual performance, as measured by tests of ability or achievement, is subject to variations according to the season of birth has been noted by many research workers. Early writers (e.g. Pintner, 1931; Pintner and Forlano, 1933) produced abundant evidence from relatively large samples of children tested in both hemispheres showing that those born in the warmer months obtained slightly higher scores on intelligence tests than those born in the colder months. Evidence was not conclusive on which season in fact yielded the highest intelligence. Fitt (1941), summarizing previous data, concluded that while the intelligence of the Autumn-born was more often found to be higher, the safest procedure was to regard the Summer/Autumn as the high-level period, with a rising tendency toward the close. Reviewing much the same previous data, however, Jones (1954) concluded that the season of birth with the highest average I.Q. was either the Spring or Summer. Most writers agreed that the difference in I.Q. between maximum and minimum was small—no more than 1 to 1½ points, on the average.

There was also disagreement on the explanation of the findings. Fitt showed that the seasonal changes also occurred with physical growth and related them to the time of conception rather than birth. He linked them with endogenous endocrine rhythms which he suggested were similar to those responsible for hibernation in other animals. Roberts (1944) eschewed such an explanation, and while producing evidence to support Fitt's findings, claimed that the reason was selective family planning—the more intelligent parents choosing the birth months of their offspring. The finding by Goodenough (1940) that seasonal differences in the frequency of births occur in the higher but not lower occupational groups supports this explanation, since children of the higher groups also have higher average I.Q. Alternative explanations have also been put forward. More recently Orme (1962), working with mentally subnormal (I.Q. 40-69) patients in hospitals confirmed the advantages of being born in the Summer or Autumn and suggested (Orme, 1963) that climatic temperature changes occurring during pregnancy might be causally related to the findings.

It is difficult to assess the real significance of these early findings. Because of the widespread belief held at the time that 'intelligence' tests yielded a measure of innate ability largely uncontaminated by environmental influences, the effect that schooling might have on the results was hardly considered. In 1959, however, John (1962) reported significantly more retarded readers with Summer birthdays than with Autumn ones, and Williams (1959) confirmed this finding and also showed that Summer birthdays occurred significantly more often with a group of educationally backward children and with another group of children referred for child guidance.

SUBSCRIPTION RATE

In the issue of *Educational Research* for June 1965, it was announced that a new Subscription rate would come into operation with the publication of the present issue.

The Annual Subscription rate for the journal is now 23/-, including postage in the United Kingdom. The price for individual issues has been fixed at 7/8d, including postage.

© *National Foundation for Educational Research in England and Wales*

Printed by King, Thorne & Stace Limited, School Road, Hove, Sussex
and published by Newnes Educational Publishing Co., Ltd., Tower House, Southampton Street, London, W.C.2.,
for the National Foundation for Educational Research in England and Wales, The Mere, Upton Park, Slough, Bucks.

DATE OF BIRTH AND SCHOLASTIC PERFORMANCE

by DOUGLAS A. PIDGEON
Deputy Director, National Foundation for Educational Research

THE fact that intellectual performance, as measured by tests of ability or achievement, is subject to variations according to the season of birth has been noted by many research workers. Early writers (e.g. Pintner, 1931; Pintner and Forlano, 1933) produced abundant evidence from relatively large samples of children tested in both hemispheres showing that those born in the warmer months obtained slightly higher scores on intelligence tests than those born in the colder months. Evidence was not conclusive on which season in fact yielded the highest intelligence. Fitt (1941), summarizing previous data, concluded that while the intelligence of the Autumn-born was more often found to be higher, the safest procedure was to regard the Summer/Autumn as the high-level period, with a rising tendency toward the close. Reviewing much the same previous data, however, Jones (1954) concluded that the season of birth with the highest average I.Q. was either the Spring or Summer. Most writers agreed that the difference in I.Q. between maximum and minimum was small—no more than 1 to 1½ points, on the average.

There was also disagreement on the explanation of the findings. Fitt showed that the seasonal changes also occurred with physical growth and related them to the time of conception rather than birth. He linked them with endogenous endocrine rhythms which he suggested were similar to those responsible for hibernation in other animals. Roberts (1944) eschewed such an explanation, and while producing evidence to support Fitt's findings, claimed that the reason was selective family planning—the more intelligent parents choosing the birth months of their offspring. The finding by Goodenough (1940) that seasonal differences in the frequency of births occur in the higher but not lower occupational groups supports this explanation, since children of the higher groups also have higher average I.Q. Alternative explanations have also been put forward. More recently Orme (1962), working with mentally subnormal (I.Q. 40-69) patients in hospitals confirmed the advantages of being born in the Summer or Autumn and suggested (Orme, 1963) that climatic temperature changes occurring during pregnancy might be causally related to the findings.

It is difficult to assess the real significance of these early findings. Because of the widespread belief held at the time that ‘ intelligence ’ tests yielded a measure of innate ability largely uncontaminated by environmental influences, the effect that schooling might have on the results was hardly considered. In 1959, however, John (1962) reported significantly more retarded readers with Summer birthdays than with Autumn ones, and Williams (1959) confirmed this finding and also showed that Summer birthdays occurred significantly more often with a group of educationally backward children and with another group of children referred for child guidance.

Further evidence

That Summer born children were also at a disadvantage in streamed schools was shown by Jinks (1964). Using data obtained from 1,315 children forming the whole of an 11 + age group in streamed schools in one County Borough, he showed that the chances of a child being in an 'A' stream varied according to his month of birth. His main results are given in Table I.

Table I

Distribution of children in different streams according to season of birth.

	Sept.-Dec.	*Jan.-Apr.*	*May-Aug.*	*Totals*
'A' stream	228	203	155	586
'B' or lower stream	201	250	278	729
Totals	429	453	433	1315

Using the marginal totals to calculate expected frequencies, a chi-square of 26.276 was obtained, which is significant at the 0.1% level.

Jinks points out that the Summer born children may have had only two years and one term in the Infant school, 'of which the one term is broken up by odd holidays, Whitsun, Open Days, Prize Days, School Plays, etc. In contrast the September intake will have had three solid years of Infant schooling behind it'. He writes: 'As a result of whatever form or system of grading the Junior school uses to stream its intake the children with most schooling behind them perform best and as a result are placed in the "A" class'.

Jinks further shows that despite the use of standardized tests with built-in age allowances, as part of the authority's 11 + procedure, the seasonal effect persists. Of 317 children who entered Grammar schools in the Borough in the year under survey, 120 had been born in Autumn (Sept.-Dec.), 103 in Spring, and only 92 in Summer.

Jackson (1964) produced further evidence confirming Jinks' findings. He showed, moreover, that as the number of streams increased (i.e. in larger schools) so the difference between Summer and Winter born children increased. In 4-stream schools almost twice as many (42% compared with 22%) Summer born as Winter born children enter the 'D' stream. Freyman (1965) confirms these findings for Junior schools, and shows further that the disadvantage of being born in the Summer months extends not only to selection at 11 + between selective and non-selective schools, but also persists within each type of secondary school; significantly more Summer born children are in the lower streams of Modern Schools than in 'A' streams, and larger proportions of the Summer born children obtain low or average marks than high marks in the yearly examinations of a selective school.

‘ Term of entry ’ and ‘ age group position ’ effects

The seasonal effect is not confined to streaming. Williams (1964), reporting on data relating to 265 children attending 8 special schools for the E.S.N., demonstrated that significantly more had Summer births and fewer Autumn births than might have been expected from the population as a whole. In seeking an explanation for the findings, Williams contrasts the ‘ term of entry ’ effect and the ‘ age group position ’ effect. The former was described by Pidgeon and Dodds (1961) who pointed out that, since there existed in this country a general (but not universal) rule that children start formal schooling at the beginning of the term in which they will become five years old, the length of Infant schooling received by any child will depend upon the month of his birthday. This is illustrated in Table II below, adapted from Williams (1964).

Table II

Effect of Birthday Month on Infant School Stay.

Birthday Month	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June	Jul.	Aug.
Enters Infants	Sept.	Sept.	Sept.	Sept.	Jan.	Jan.	Jan.	Jan.	Apr.	Apr.	Apr.	Apr.
No. of Terms	9	9	9	9	8	8	8	8	7	7	7	7

It is not an unreasonable assumption that children who have received more schooling will have learned more, and hence, when examined, will produce better performances. The children with longer schooling, however, will, in any given school year group, be the older ones, and Pidgeon and Dodds, in a study involving some 1,600 children in six large Junior schools, obtained a correlation of .979 between age and length of schooling. Williams points out that it is not necessary to assume that a ‘ term of entry ’ effect is operating alone and offers the alternative ‘ age group position ’ effect as an explanation. Summer born children, he argues, usually retain their position of being youngest in their age group throughout the Junior school, and this imposes additional stress since teaching levels will tend to be harder for younger children. In an attempt to distinguish between these two effects, he divided his group according to their month of birth *within* the three terms. His analysis of his data however failed to distinguish between the effects, though it suggested that the age group position effect might be the more powerful of the two. He concluded that further investigations were necessary.

Evidence that the ‘ age group position ’ effect does indeed operate can be obtained from Sweden, where children born during any one calendar year start school together at the beginning of September. Johansson (1965) showed that both assessments and tests of school readiness gave low but significant correlations with age at the start of school. Ljung (1965) produced evidence to show that of children in Grades 4, 6 and 8 who had repeated a grade at some stage, significantly more had birthdays in the last months of the year.

Studies by Anderson (1964) and Postlethwaite (1965) indicate that (within certain limits) the length of previous schooling has no significant effect on standards achieved by the age of 13. In other words, it makes absolutely no difference to their eventual achievement whether pupils begin school at 5, 6 or 7. There are various factors, such as the amount of pre-schooling, which might account for this. Nevertheless, it does suggest that differences in the achievement of Autumn and Summer born children are attributable to other causes than the longer period at school enjoyed by the older children under our present entry system; if a two-year difference between children of various countries can be ironed out by the time they are 13, a difference of two terms should not affect achievement.

The situation in ' streamed ' schools

One further piece of evidence needs to be given. In the study referred to above, Pidgeon and Dodds demonstrated that if due allowance for age was made on Reading Test performance, *in schools which streamed for age*, a residual effect due to length of schooling occurs in the 3rd and 4th junior school years, although no effect was present in the 1st and 2nd years, or in any year of junior schools which either streamed by ability or grouped children randomly into classes within the year group. By way of explanation they suggested that, in these special circumstances of grouping by age, the oldest ' stream ', being older and having more schooling, will have reached a higher standard and, by giving the appearance of being ' better ', will therefore be advanced more rapidly than might otherwise have been the case, while with the youngest ' stream ' the opposite will occur. Since this effect is cumulative it is only noticeable during the latter years of the junior school. While this explanation fits the facts, Pidgeon and Dodds point out the need for confirmation from further research, especially as an increasing number of schools are changing over from ability streaming to grouping by age.

Conclusions

While the possibility may exist that children born during the warmer summer months obtain slightly higher scores on ' intelligence ' tests, there is little doubt that the advantages gained by these children are small compared with the disadvantages conferred on them by being born at the ' wrong ' end of the school year. The evidence from research demonstrates clearly that the younger children in any school year group are at a disadvantage compared with the older children. They tend to be in the lower streams of both primary and secondary schools and, despite the use of age allowances in standardized tests, appear to suffer slightly in 11+ examinations. In addition, they tend to predominate in Backward groups and remedial classes.

The disadvantage of being the youngest pupils in a class is certainly not helped by an education system that offers these same younger pupils rather less schooling than their older schoolmates. At the same time it is clear that, although having one point of entry only into an education system and thus giving all children the same length of schooling might help, it will not, by any means, entirely remove the disadvantage.

The cause, if such a word can be used, of this phenomenon, is the fact that children are born at different times of the year, and it is hardly conceivable that this can be changed. The problem, therefore, is to remove the effect. Starting school at one time only in the year will probably lessen, but not remove, the effect. The abolition of streaming will certainly remove one very obvious and distasteful effect, but it will not eradicate the problem—and while classes consist of younger and older children, the younger may continue to suffer in other ways.

One solution which suggests itself is age grouping; that is, putting into teaching units children separated by only a month or two of age. The difficulty here is that such a practice is only possible in fairly large schools, and even where it is practised the kind of results obtained by Pidgeon and Dodds may be found. Presumably such a result, where older children do better than expected for their age and younger children do worse, must largely be due to teacher attitudes, and to remove this effect it would be necessary to change these attitudes. Granted this can be done, then for large schools at least, age streaming, coupled perhaps with one point of entry into the education system, might help to remove the disadvantage of being born in the summer.

It might seem, however, that the only true solution to the problem is the individualizing of learning, in which children compete only with themselves, and are not taught as a class. In these circumstances the younger child (of equal ability as an older one) would merely be at an earlier stage of development at any given point in time. The use of mixed age groups—that is, class groups formed from pupils of all ages such as may now be found in small one-teacher schools—would probably help to achieve this individualization, but it must be recognized that it also implies a change of approach from present class teaching methods.

Bibliography

Anderson, I. H. (1964). 'Comparisons of the Reading and Spelling Achievement and Quality of Handwriting of Groups of English, Scottish, and American Children'. Co-operative Research Project, No. 1903, The University of Michigan, U.S.A.

Fitt, A. B. (1941). *Seasonal Influence on Growth, Function and Inheritance.* New Zealand Council for Educational Research.

Freyman, R. (1965). 'Further Evidence on the Effect of Date of Birth on Subsequent School Performance, *Educational Research*, VIII, 1.

Goodenough, Florence L. (1940). 'Can we influence mental growth?', *Ed. Record*, Supplement No. 13.

Jackson, B. (1964). *Streaming: An Education System in Miniature.* London: Routledge and Kegan Paul.

Jinks, P. C. (1964). 'An investigation into the effect of date of birth on subsequent school performance', *Educational Research*, VI, 3, 220-225.

Johansson, B. A. Criteria of School Readiness, Factor Structure, Predictive Value and Environmental Influences, Stockholm Studies in Educational Psychology 9, Almqvist & Wiksell, Stockholm.

John, E. (1962). 'The Age Factor in Reading Retardation, Researches and Studies', University of Leeds Institute of Education, No. 24, pp. 1-7.

Jones, H. E. (1954). 'The Environment and Mental Development', *in:* Carmichael, L., ed., *Manual of Child Psychology* (2nd edition). New York: Wiley; London: Chapman & Hall, pp. 631-696.

Ljung, B. O. (1965). 'The Adolescent Spurt in Mental Growth', Stockholm Studies in Educational Psychology 8, Almqvist & Wiksell, Stockholm.

Orme, J. E. (1963). 'Intelligence, season of birth and climatic temperature', *B.J.P.*, vol. 54, part 3.

Pidgeon, D. A. and Dodds, E. M. (1961). 'Length of schooling and its effect on performance in the junior school', *Educational Research*, vol. III, no. 3, 214-221.

Pintner, R. (1931). 'Intelligence and month of birth', *J. App. Psy.*, vol. 15.

Pintner, R. and Forlano, G. (1933). 'The influence of the month of birth on intelligence quotients', *J. Ed. Psy.*, No. 24.

Postlethwaite, N. 'Organisational Differentiation and School Achievement, A Preliminary Study of Mathematics Performance in Different School Systems'. Unpublished Fil.Lic. thesis, University of Stockholm.

Roberts, Fraser J. A. (1964). 'Intelligence and the season of conception', *B.M.J.*, vol. 1.

Williams, P. (1959). 'The Growth of Reading Vocabulary'. Unpublished Ph.D. thesis, University of London.

Williams, P. (1964). 'Date of birth, backwardness and educational organisation', *B. J. Ed. P.*, vol. XXXIV, part 3.

A SURVEY OF INSTRUCTIONAL TELEVISION RESEARCHES

by H. Barrington,

Senior Lecturer in Education, Edge Hill College

'IT now seems clear, however, that television offers the greatest opportunity for the advancement of education since the introduction of printing by movable type. This comparison is made soberly'. This opinion was voiced by Dr. J. C. Poolock of New York University in 1960.

If Dr. Poolock was correct in his conclusion then it would appear that, in Britain in 1965, we have not yet taken full advantage of the opportunity offered by television for the advancement of education.

One of the reasons why television has not been more frequently used for teaching has been that many teachers have a negative attitude to instructional television. It would appear reasonable to assume that this negative attitude stems from a belief in the value of conventional methods of teaching rather than from proof of the ineffectiveness of television for teaching. Objective evidence of the value of instructional television has not been readily available to the classroom teacher. A former Minister of Education, Sir Edward Boyle, replying to a parliamentary question on 2 May 1963, said: 'I know of no published research in this country, based on controlled experiments, into the efficiency of television teaching in schools and I am considering now whether it might be valuable to initiate research on its effects'.

Several investigators have made the point that information about the effectiveness of television for teaching would influence the acceptance of instructional television by teachers; as, for example, the statement by Carpenter (1960): 'The acceptability of television by teachers depends upon its proven and demonstrated effectiveness'. Dr. Wigren of the National Educational Association, Washington D.C., reinforced this argument when he remarked: 'What they (teachers) resist are unwarranted claims, which cannot be supported, for television's use in the classroom'.

What evidence is there of the effectiveness of television for teaching? This is one of the questions which this survey should help to answer.

1. The effectiveness of instructional television

Where television has been used in the classroom there has been, in general, a short period of preparation before the broadcast and a follow-up period after the broadcast. No general figure can be given for the time devoted to these three divisions, but in a nation-wide project involving some 200 public school systems in the United States of America, elementary schools gave an average of 5 minutes preparation for a 20-minute broadcast. In secondary schools the average time given to the follow-up of a 30-minute broadcast was 25 minutes. No general figure could be given for either follow-up time in elementary schools or preparation time in secondary schools.

Several research workers have reported no significant differences between the achievement of pupils taught by television and pupils taught by conventional methods. Anderson and Vandermeer (1955), found that television instruction was as effective as personal instruction for teaching the use of the slide-rule. Kanner, Runyon, and Desiderato (1955), obtained similar findings using television for training army personnel. Tannenbaum (1956), found that the achievement of pupils taught biology by television was better than that of pupils taught by the teacher in person, although the difference was not significant.

The Cincinnati Board of Education (1958), reported that a comparative study of the teaching of high school chemistry had indicated no significant differences between the results obtained by pupils taught by television and those taught by conventional methods. Similar findings were reported by Seibert (1958), after experiments with the teaching of physics.

Since 1954, Carpenter and Greenhill, financed by the Fund for the Advancement of Education, have conducted research with television for teaching at Pennsylvania State University. Their report no. 1 (1955) reported no significant differences between the achievement of pupils taught by television and those taught by other methods. Report no. 2 (1958) reported similar findings. Greenhill (1959), said that 500 students who had studied zoology by television lessons had obtained good results. He reported that experiments with television over a five-year period indicated that television instruction was effective.

Williams (1954), described a comparative study of the effectiveness of teaching by television carried out in Toronto, Canada. One hundred and eight high school students were randomly allocated to 4 teaching groups. An unfamiliar topic was chosen: then, one group studied it by television, a second group read the script of the broadcast, the third was present in the studio during the broadcast, and the fourth heard only the soundtrack of the broadcast. After testing, the students who had viewed the programme proved to have learned significantly more than all the others.

Heminghaus (1957), noted that the results of a study conducted in St. Louis schools had indicated that the achievement of television-taught pupils was equal to that of conventionally taught pupils.

Findings of no significant differences were reported by Dowell (1956), who taught electronics by television, Shimberg (1957), who used television for teaching nursing, and the Signal School, Fort Monmouth (1957), where television was used for teaching radio mechanics.

The United States Naval Training Devices Centre (1957), provided a summary of findings of researches into teaching by television up to 1956. Conclusions formed were that television teaching was as effective as conventional teaching: the teaching was well liked, subject matter was retained, and factual programmes were the most successful.

Ambrosino, Dunstam, and Haake (1957), reporting experiments with television in Schenectady high schools, found the learning of television and control groups to be equal. Dreher and Beatty (1960), reported similar findings after investigations in San Francisco. A comparative study of the teaching of physics by television and other methods was reported by Boeck and Washton (1961). One group of pupils received instruction by television, a second group received the television lesson and had their own teacher in the room to develop the lesson, and a third group had a conventional lesson given by this same teacher. No significant differences were found between the achievements of the groups.

General conclusions, based on a review of investigations conducted in the United States, were made by Seldes (1957). He claimed that the value of television as an instrument of direct teaching had been demonstrated and that—though the evidence, at that time, was not conclusive—no negative evidence had been adduced.

One investigation in Britain, which appears to have obtained reliable data on which to base conclusions, was conducted by the Institute of Education of Nottingham University during 1958. Daniels (1959), reported that four groups of secondary modern school pupils studied a science topic by different methods. Of the four groups involved, one viewed a television lesson, one heard only the sound track of the broadcast, and the third had a conventional science lesson. The fourth group received no instruction, and was used as a control group. The conclusions were that some teaching was better than no teaching at all but that there were no significant differences to indicate that any of the three methods was superior. This experiment was probably the only controlled experiment of its type in Britain at this time and for some time afterwards.

Barrington (1961) carried out an investigation into the effectiveness of television for teaching physical science as compared with conventional instruction in the laboratory. The experiment was intended as a pilot study for a later more detailed investigation. Two groups of secondary modern school boys of average ability, aged between thirteen and fourteen years of age, formed the sample. One group followed three television lessons from the B.B.C. and the second group had conventional lessons in the laboratory. The results indicated that the laboratory group had done a little better than the television group but that the difference was not significant.

A writer in the *Review of Educational Research* (no. 4, 1960, p. 354), concluded that, despite the fact that the majority of findings showed no significant differences in favour of either television or conventional teaching, there was certainly no evidence for claiming that television was inferior to personal teaching.

Tanner (1961), in a survey of 281 investigations, noted that 90% of comparisons between television-taught and conventionally-taught pupils resulted in findings of no significant differences in achievement. He suggested that there *should* be differences and that the tests used in the investigations might not have been sensitive enough to measure the differences which existed. He pointed out that most studies had been concerned only with achievement and that no attempts had been made to measure intangibles in the learning situation.

Skornia (1960) gave a similar verdict: ' There must be differences. But the measuring techniques and tools available, designed only to measure print based effects, or effects which can be communicated in writing, are inadequate to measure the kinds of learning and conditioning which are now occurring '. Again, Erickson (1960), suggested that failure to reveal significant differences may be due to the inability of available devices to discriminate differences due to method. Williams (1962), made the same observation and Holmes (1962), stated: ' We must not be satisfied with results of "no significant differences" for there must be differences '.

However, Kumata (1960) claimed that the few experiments expressly designed to measure some of the apparent intangibles of the learning situation, such as the development of critical thinking, also indicated no differences between any learning changes in television-taught classes and in conventional classes.

Purdue University Report no. 9 (1959) claimed that television lessons had proved inferior to conventional lessons in mechanical engineering courses, but in most other cases no significant differences had been found. The final report on experiments conducted at Miami University (1960), concluded that there was little loss on television-taught courses as compared with conventional methods of instruction.

In 1957, the Ford Foundation and its Fund for the Advancement of Education (1961), sponsored ' The National Programme on the Use of Television in the Public Schools '. Two committees of experts, on education and testing methods, were set up to evaluate the results of comparisons between television-taught and non-television-taught pupils in elementary and high schools.

During the first two years, from 1957 to 1959, 251 comparisons were made of the achievements of television-taught pupils with those of control pupils. One hundred and sixty-five comparisons favoured television and 86 the control pupils. Ninety of the comparisons were significant and 161 not significant. Of the comparisons showing significant differences, 69 favoured the television groups and 21 favoured the control groups. These results were obtained from comparisons where the groups had been equated on scholastic aptitude and pre-test score or where such initial differences between groups had been held constant so that legitimate comparisons could be made.

Instructional television has been used on a large scale in Hagerstown, Washington County, Maryland. Weiss (1957), reported that teachers, pupils, and administrators were enthusiastic about the quality of learning and the greater subject-coverage made possible by the use of television. The results of standardized tests showed television classes to be out-performing conventional classes. Cassirer (1960), noted that where other methods of teaching arithmetic had failed, television lessons in Hagerstown elementary schools had produced, in a 7½-month period, a mean gain in attainment of 17 months in 18 classrooms. Brish (1960) reported that the Stanford Achievement Test in science had been given first in September, 1957, and again in May, 1958, to about 900 sixth-grade children studying science by television, and to 600 having conventional instruction. At all levels of ability the television pupils made greater gains than the conventionally-taught pupils. For the pupils of below average ability the increased attainment in months indicated by the scores of the achievement tests was double that of the conventionally taught pupils.

The Washington County Closed Circuit-Television Report (1964) reported the results of the experiment with television for teaching in Hagerstown during the five-year period 1956 to 1961. The report concluded: ‘ The use of television can bring substantially improved achievement. In many courses taught with the aid of television the increase is statistically significant. Achievement can improve year by year with television long after any novelty effect has worn off ’.

Maeda (1962), described an experiment with television lessons in a remote village school at Doruba in Japan. After six months of viewing a series of science lessons, achievement tests, given by experts, indicated that the ability of the pupils had been remarkably raised. After one year, Ministry of Education testers found the pupils’ ability in science to be higher than the national average.

Jacobs and Grate (1962), reported the results of a pilot study of the use of television for teaching sixth-grade science in Cincinnati schools. The more able children did better using television, although not significantly so, and the less able pupils did significantly better by conventional instruction.

Deutschmann, Barrow, and McMillan (1962) reported a comparative study carried out at Michigan State University in 1961. One group of students saw a film, a second group viewed the film over television, a third group had a science lesson in a laboratory, and a fourth group had a laboratory lesson and saw the film. An assessment was then made of the relative efficiency of the different methods of instruction. Efficiency was defined as ‘ the discrepancy between the amount of relevant information and the amount of irrelevant information a student learned ’. Analysis of variance produced an F value which was highly significant ($P<.001$). The order of efficiency of methods was: 1. Television; 2. Film; 3. Film and Laboratory; 4. Laboratory.

A similar investigation was carried out by Harris (1964), who compared the achievements of groups taught by the following methods: Laboratory, Laboratory and Colour Film, Colour Film, Television and the Film (black and white). Television and Film proved the most efficient although the differences were not significant.

Royds (1963), carried out an investigation by questionnaire of televised science lessons. His sample consisted of 1,578 children from a wide area of Britain. Of the sample, 872 were in primary schools, 649 in secondary modern schools, and 57 were in grammar schools. His main conclusions were that learning from television depends on the ability and the environment of the student and that the method of presentation of the lesson is very important.

Barrington (1964), compared the achievement of pupils who had studied physics by television with that of pupils taught by conventional methods. His sample consisted of five groups of thirty pupils taken from the first and second years of a secondary modern school. One group was taught by television lessons integrated with laboratory lessons. A second group had conventional laboratory lessons. A third group took the television lessons in a classroom and a fourth group had conventional classroom lessons based on a textbook. A fifth group was used as a control group. No significant difference was found between the achievement of the T.V./laboratory group and the conventional laboratory group. The achievement of the T.V./classroom group was found to be significantly higher ($P < .05$) than that of the conventional classroom group.

Schramm (1962) reported that of 56 comparisons of television science instruction with conventional instruction, it was found, on testing for significant differences at the .05 level or higher, that 20 studies favoured television, that 7 studies favoured conventional instruction, and that in 29 studies there were no significant differences. At college level, 28 comparisons resulted in one study in favour of television, one in favour of conventional teaching and 26 reporting no significant differences. Schramm concludes by saying: ' There can no longer be any doubt that students learn efficiently from instructional television '.

2. The development of instructional television

Development in the U.S.A. The first experiments with television for teaching appear to have taken place in the United States of America in 1950. Siepmann (1952) noted that interim conclusions from studies carried out at Montclair State Teachers' College were that television was effective for demonstration teaching and that retention of information was satisfactory. In 1954, lessons, prepared under the direction of lecturers at the College, were televised to schools in New Jersey.

Since that time, many investigations into the effectiveness of television for teaching have been carried out in the United States and the acceptance of the findings from these investigations has resulted in the large scale use of instructional television in that country. The National Programme, which employed broadcast television and the Hagerstown system, which used closed-circuit television, have already been mentioned in this survey. The former involved nearly 200,000 pupils in some 200 public school systems and the latter 18,900 pupils in 48 schools. These 48 schools were equipped with 700 receiving sets and 25 lessons were televised each day.

Perry and Beredy (1960), Voorhies (1961), and the Ford Foundation (1961), all reported on the programme of the Midwest Council on Airborn Television Instruction. Broadcasts were transmitted from an aircraft flying at 23,000 ft. above Indiana to a circular area of some 200 miles radius from the aircraft. The service could provide 72 televised lessons each day for about 5,000,000 students.

Baxter (1960), noted that the National Broadcasting company's science programme ' Continental Classroom ' had an estimated weekly audience of 1,700,000 students and that 200 colleges were giving credits for television chemistry and physics.

Scanlon (1960), reported that the University of North Carolina television station was providing television lessons for 250,000 pupils in 400 schools. He also noted that at Anaheim, California, 12 schools had been linked by cable so that 3,000 pupils could use television lessons.

The Ford Foundation Report (1961) reported that three educational television stations in Alabama could provide a daily state-wide service and that at that time 300,000 pupils in 600 schools were using televised lessons.

Cassirer (1960), reported the extensive use of instructional television in the schools of Philadelohia where the first broadcasts had started as early as 1952.

The Independent Television Authority's Report (1961) said that, in 1961, 3,000,000 students in 7,500 schools in America were receiving daily teaching by television and that 250 colleges were giving credits for televised courses. It was also noted that in 1961 51 educational television stations, 6 commercially owned educational television stations and 250 closed-circuit television systems were operating in America.

Experiment and development in Britain. The earliest investigation of the use of television for teaching in Britain appears to have taken place in 1950. Brown (1954), described experiments with the normal B.B.C. service and children in a London school. His investigations ceased in 1951 before reliable conclusions had been formed but he did say that children could assimilate a remarkable amount of information from a television programme received under classroom conditions.

The B.B.C. conducted a pilot experiment in six London schools in 1952. Love (1963) stated that this was the only experiment conducted by the B.B.C. into teaching by television before 1957 when the B.B.C.'s School Television Service commenced. No controlled experiments appear to have been conducted after the start of the B.B.C. service. No valid conclusions were claimed by the B.B.C. after their investigation.

The School Broadcasting Council Report (1959), stated that a non-specialist teacher appeared to have obtained satisfactory results by using televised science lessons. The report also claimed that television lessons could be useful to the specialist teacher of science.

Roy (1959) claimed that pupils obtained useful background information from television lessons but that attempts at direct teaching were not always successful.

None of these claims resulted from carefully controlled studies.

In 1959 a television research unit, sponsored by Granada T.V., was set up at Leeds University under the direction of the late Dr. J. Trenamen. This unit however was concerned with the social rather than the educational aspects of television.

Some universities and education committees, sometimes with the help of television companies, have conducted experiments with television for teaching. Few of these could be described as controlled investigations designed to evaluate instructional television as compared with conventional teaching methods. Apart from the findings of the few investigations mentioned in section 1 of this survey, the evidence concerning the effectiveness of instructional television in Britain has been gathered from the subjective opinions of teachers.

In September, 1961, the I.T.A. companies co-operated to provide a national service for schools additional to the existing B.B.C. service. Associated Rediffusion reported in 1962 on the first four years of the service for schools, and it was stated: 'This is not a statistical survey. The findings of the report, though neither exhaustive nor conclusive, are based on considered judgements'. How valid these may be, is shown by the following study.

Daines and Neilson (1963) investigated the reliability of the judgements of teachers regarding television lessons. In this experiment, a filmed record of fourteen year old girls of average to good intelligence viewing a series of television lessons, was studied by 29 teachers. The teachers were asked to make written judgements of the level of interest shown by the class. After analysing the judgements the investigators remarked: ' The query must project itself whether, in this sort of work, the interest observed is in the eye of the observer rather than that which is observed '. They also remarked: ' The reactions of teachers to their children's behaviour would seem to have no consistent pattern. This we find a little disturbing as to some considerable degree the television authorities tend to rely upon just such judgements as these '.

Various experiments have been conducted with forms of closed-circuit television. For example, in 1961, science lessons given in Hayes Grammar School were transmitted to Barnhill Secondary School. The Cambridgeshire Education Authority linked three of the village colleges by closed-circuit television. In 1963, Hampshire Education Committee set up a studio in Warblington Secondary School and lessons were transmitted to two other secondary schools. A closed-circuit link was also set up between Avery Hill Training College and Kidbrooke Comprehensive School. In the majority of these cases (the Warblington experiment being a notable exception) no attempts were made to assess rigorously the effectiveness of television for teaching and no quantitative findings were published.

Some reports in the press of the development of closed-circuit television for education have been misleading. Services assumed to be operational have been found, on investigation, to be only drawings on paper or ideas in somebody's head. Several cities, including Hull and Liverpool, have plans for local educational television services. In April 1963, television lessons were broadcast from Glasgow College of Commerce to a school some five miles away. The broadcast was claimed to have been a success both technically and pedagogically. The Chairman of Glasgow Education Committee (1963), said that, as a result, a Glasgow Schools Television service would be developed.

There have been demonstrations of television installations for teaching in medical schools, notably at Guy's Hospital, but they have been demonstrations rather than controlled experiments. Comparative studies of the effectiveness of colour television for medical teaching are to be carried out at Guy's.

An interesting experiment is described in a report of A.B.C. Television (1963). Cambridge University Department of Education and A.B.C. Television conducted an experiment to measure the amount of information gained from a television programme. Although the results were not published in the report, it was claimed that the method has worked successfully.

A survey of the application of television to education in Britain was published by the Central Office of Information (1963). It was neither detailed nor exhaustive but it was a welcome addition to the literature in a field of research where pertinent information was, and still is, in short supply.

Developments in Europe and the U.S.S.R. Puglisi (1962), reported that in Italy a direct teaching service for schools, Telescuola, commenced in 1958. Pupils following the lessons worked with a co-ordinator who maintained contact with the studio in Rome by sending in reports, attending meetings in Rome, and using guides issued for co-ordinators. The pupils used special text books prepared by the television teachers and a monthly magazine sent to all pupils. Each week the best pupils were able to qualify for a visit to Rome to meet their teachers. Amaldi (1960), claimed that Telescuola pupils had succeeded in solving problems which many pupils attending secondary schools could not cope with.

The French television service to schools commenced in 1951. Travers (1960), reported the introduction, in 1955 at St. Cloud Teacher Training Centre, of closed-circuit television for training teachers in the use of television for teaching. Guillin (1960), said that science programmes were broadcast for secondary pupils and aimed to provide material difficult for the science teacher to obtain. He noted also that a direct teaching service was provided for pupils unable to attend school. Systematic studies of the value of television for teaching have been instigated at the Lycée de Sèvres.

Dieuzeide (1960), noted that experiments had taken place in Belgium, Netherlands and Switzerland but that the results were inconclusive.

Sunesen (1960), reported a negative attitude to educational television in Denmark and that there were no definite plans for a school television service.

Ogren (1960), reported that the first instructional television experiments in Sweden took place in 1951. They were mainly concerned with testing equipment. A further investigation in 1957 was conducted by questionnaire: the replies from 100 schools indicated great promise for a schools service and two committees were formed to investigate television for schools and for adult education.

Bohlen (1960), said that a schools television service appeared to be out of the question for Germany.

Suchodolski (1960), summed up the position of instructional television in Poland as: ' Television is still very little known and teachers have no experience in this field '.

Matthews (1963), said: ' School television in the U.S.S.R. is still in its early infancy '. He reported that experiments were being conducted in Odessa with the teaching of physics and biology by television and that the prospects for the development of instructional television seemed good.

Australia and Canada. Bull (1960), reported the start of general television in Australia in late 1956 in Sydney and Melbourne. From the beginning of the service, experimental programmes were provided for schools. Some of the programmes were for pre-school children. The Australian Broadcasting Company's programmes for schools were available in 1960 to schools within a radius of 100 miles of the capital city of each state. Some experiments with instructional television commenced in 1962 at Melbourne University.

Williams (1954), described the first serious experiment with instructional television in Canada. This took place in Toronto and has been referred to in section 1 of this survey. Cassirer (1960) noted experiments with televised lessons in Halifax schools. These experiments had indicated that television could not replace good teachers but that it could help to mitigate staff shortages.

Progress in Japan. Hatano (1960) and Maeda (1962), reviewed the development of television for teaching in Japan. In February, 1953, only fifteen months after the start of the general television service, the Japanese Broadcasting Company, N.H.K., commenced television broadcasting for schools. An educational channel was provided in 1958 and educational programmes were then broadcast on both channels. By 1962, The National Federation of Study Groups of Radio and Television Education had been formed by teachers using the television lessons and 80% of all primary and secondary schools were members of the federation. In 1963, special programmes of G.C.E. type work were provided for children unable to obtain normal secondary education. Those who follow these courses can attend school at weekends and during holiday periods and obtain help from teachers. Arrangements are made for the students to eat and sleep in the schools. At the end of the course, the students take the same examination as pupils in full-time attendance at school.

An International centre. Wilson (1962), reported the establishment, in London, of the Centre for Educational Television Overseas to inform, advise, train, and provide materials for, people from any country wishing to use television for education. Many overseas countries, notably in Africa, see in television a method of providing some education for many people who would not otherwise receive any. They are not interested in the comparative effectiveness of instructional television.

Television in Africa and India. In Uganda a public service, helped by advertising, provided programmes for schools in early 1964. A series aimed at preparing pupils for the G.C.E. examination began in November 1962 on Kenya Television. The Ghana government intends to develop extensive television coverage of the country with educational television a major concern. Nigeria has a very limited amount of educational television whereas Southern Rhodesia has separate educational and general television services. In Northern Rhodesia the education departments of the copper companies, together with the government, were developing a television service.

However in many African countries the costs involved in developing an extensive television service are high and cannot be afforded. Some attempts have been made to develop television on a commercial basis but, owing to the small number of people with receiving sets, the audience is limited and hence advertisers are few in number.

The position is little better in India. In early 1962, All India Radio, financed by grants from the Ford Foundation, broadcast television lessons to Delhi schools. A National Institute of Audio Visual Education has been founded and experiments are conducted with closed-circuit television.

3. Aspects of instructional television receiving special attention

The ability of television to hold the attention of pupils has been noted by several workers. Bull (1960), said that Australian experiments had indicated that television had a capacity to arrest the attention of pupils. Brish (1961), claimed that the television screen gave the teacher an effective means of directing and controlling the attention of pupils. Barrington (1964), also reported similar findings.

Various suggestions have been made as to why television should have this effect. Vernon (1953) in the early days of television stated that: ' The impression of reality created by some types of television material is exceedingly valuable in attracting and interesting the viewer. It seems to be the main factor in stimulating him to attend to the programme and to follow it ', and Guillin (1960) claimed that television gave an element of suspense which film did not.

It has been suggested that pupils may concentrate on a television lesson because they are aware that if they miss something they will not be able to turn back the programme. This idea was supported by Hemminghaus (1957).

' Eye contact ' between the television teacher and the pupils has also been noted as a possible reason for the attention holding property of television. Meaney (1962) said: ' Television is, paradoxically, a kind of one to one presentation '. An A.V.C.R. abstract (1961), described an experiment to determine if eye contact with the viewers could be associated with (a) greater attention; (b) greater interest; (c) greater information gain. The results indicated that the amount of eye contact time made no significant difference in any part of the investigation.

Daines and Neilsen (1963) reported: ' There appears to be considerable evidence for what can only be called tele-hypnosis on the part of the children '. As a television picture is produced by intermittent illumination, its effect on the viewer—especially one seated near to the screen—could be similar to that of intermittent photic stimulation used in medical practice for producing epileptic discharges. Pantelakis, Bower and Jones (1962) reported convulsions in children which were caused in television viewing. They further reported an incidence of television-induced epilepsy among epileptic children of 5.6%.

Hayman (1963) studied the effect on attention of the position of the pupil in relation to the television screen. Tests were given to determine the achievement of pupils in (a) understanding spoken Spanish, and (b) the pronunciation of Spanish. The room position of each pupil was marked by the teacher on the pupil's answer-paper. It was concluded that visual acuity was almost certainly more important to the speaking skill than to the understanding skill and that the effect on learning of the viewer's position depends on the learning task. If the visual aspect is very important, then the position of the viewer is vital.

The contributions made by sound and vision to the learning situation were also investigated by Trenamen (1959), although he did not concern himself with the pupil's position in the room. He claimed that the spoken part of the television message is more important than the visual. The one visual factor satisfactorily assessed was that of visual movement and he found that, surprisingly, visual movement failed to contribute to comprehension. The most important factor for comprehension was concreteness which he described as ' the extent to which the programme is focussed on what can be perceived'.

The effect on the attention of pupils of the length of a television lesson was investigated by Bridges (1960). His findings indicated that a ' fall off ' of attention most frequently occurred after 25 minutes.

Special consideration has been given to the effectiveness of television with pupils at different ability levels. Kanner, Runyon, and Desiderato (1956) claimed that television was more effective than traditional methods for teaching pupils of lower aptitudes. In the Hagerstown experiment, pupils of I.Q. 90 and below made a gain of 14 months in arithmetic attainment over a period of 7½ months.

Jacobs and Grate (1962), however, seem to imply the opposite when they say: ' There appears to be a trend for the conventional method to become more superior to the television method of instruction as the ability level decreases'. Boeck and Washton (1961), also reported that television instruction was better than conventional instruction for the average pupil only. Pupils of I.Q. 120 and above and I.Q. 100 and below, did better with conventional instruction.

The situation regarding the effectiveness of television for teaching at the different levels of ability was perhaps best expressed by Schramm (1962) whose judgement was that: 'It must be admitted that we do not yet understand the relation of mental ability to differential learning from television'.

Devitt (1963), investigated the effectiveness of television lessons in providing intellectual stimulation for gifted pupils. He concluded that television instruction had a desirable effect on gifted pupils and that subject matter assistance was less important for learning from television than students' awareness of the value placed by their teachers upon the programmes presented.

The relationship between the level of interest of the pupil in a television lesson and his subsequent achievement was investigated by Becker (1964). Three methods were employed to determine the level of interest of students taking television instruction. One method attempted was the measurement of galvanic skin responses. No significant correlation was found between knowledge gain and level of interest. Though overt responses indicated less interest when the teacher appeared on the screen, measurement of galvanic skin responses indicated at this time a state of alertness and tension which should be conducive to learning.

A frequent criticism of the use of television for teaching has been that pupils cannot make direct contact with the television teacher. Mercer and Becker (1955) and Hemminghaus (1957) noted this criticism. However, Carpenter (1958) reported that two-way communication between studio and classroom, had not been shown to improve results when compared with television lessons without two-way communication. Kanner (1957), also claimed that a lack of two-way communication was not detrimental to learning by television lessons and Cassirer (1960), contended that experience with television had cast some doubt on the assumption that opportunities for asking questions and for classroom discussion are most valuable for fostering learning.

The necessity for such communication was commented upon by Erickson (1960) in the following terms: 'Electronic two-way communication with students, often thought to be necessary in the early development of an installation, seems to dwindle in importance as a body of experience is acquired'.

Wolgamuth (1962) investigated the comparative effectiveness of two-way communication in television teaching. Eighty college students were divided into four groups each of which followed the same subject matter by different methods. Two of the groups had two-way facilities. The results indicated no significant differences between the groups in achievement or retention. No consistent evidence was found of the methods producing any significant effect on the attitude of students to instructional television.

Kumata (1960) claimed that although students would like two-way communication, its absence did not affect achievement. Tanner (1961), was certain that two-way communication was not important to television teaching in his view: 'Two-way communication may impede the quantity of information to be disseminated during a given period'.

Follow-up procedures after televised lessons have received attention. A Ford

Foundation Report (1961) noted that in the St. Louis experiment of 1955 the achievement of the television pupils in science was equal to that of the conventionally-taught pupils and it was suggested that the best results from television lessons would be obtained from a combination of the televised programme and follow-up work by the teacher in the classroom.

Rehage (1959), reported conclusions from experiments in Pittsburgh with television for teaching. It was suggested that the effectiveness of televised lessons depended to an important degree on the quality of the follow-up work in the classroom.

Perry (1960) claimed that television was neutral and that its value must therefore be in proportion to the skill with which it is used—the skill in getting lessons on the air and the skill in follow-up in the classroom. Brish (1961), reporting results in Hagerstown, stressed the importance of combining televised instruction with work in the classroom.

Kanner and Marshall (1963), carried out an investigation to evaluate attempts at using television for improving the learning and achievement of army personnel. Eight companies of 150 men of similar ability formed the sample. One company received television instruction including a review/preview technique when trainees were able to see in the evening a review of their lesson and a preview of the next stage of training. This technique did not include the specifics of classroom instruction. Results indicated that the television-taught company performed significantly better than all other companies and that both high and low aptitude trainees taught by television learned more than their counterparts receiving conventional instruction. The ' after hours ' review/preview was believed to be the major reason for the success of the television-taught trainees.

Follow-up procedures were also investigated by Hayman and Johnson (1963). A second viewing of a television Spanish lesson (i.e. exact repetition) was evaluated under three conditions:

(*a*) Compared with a single viewing;

(*b*) Compared with a single viewing and teacher directed practice (varied repetition); and

(*c*) As an addition to a single viewing and classroom practice compared with the same conditions without repeat viewing.

It was found that exact repetition significantly increased learning where no other follow-up was used. But it was not as effective as teacher-directed follow-up *providing* such follow-up is ' skilfully presented '. Not *all* variations will be better than exact repetition.

4. Research into long-term retention of television learning

Postman (1961) said: ' Perhaps the most important single criterion in the evaluation of training procedures is that of long-term retention '. However comparatively little research has been conducted in the field of long-term retention of television learning.

Anderson and Vandermeer (1955) said that their television taught pupils forgot what they had learned, about the operation of the slide-rule, more quickly than did their conventionally-taught pupils.

Kanner, Runyon, and Desiderato (1955) found no differences in retention between television and conventionally taught classes. Seldes (1957), also reported that there appeared to be no differences in retention between television and conventionally taught pupils but he said that the evidence available was inconclusive. In the Toronto experiment, Williams (1954) noted that tests given eight months after the teaching period, had indicated that the relative levels of achievement between television taught pupils and others had been retained. He claimed that forgetting was unrelated to the medium of instruction.

Jacobs and Bollenbacher (1960) gave achievement tests to 180 television-taught pupils and 180 conventionally-taught pupils twenty-two months after the end of the teaching period of the experiment. The results indicated that the retention of subject matter was similar for both groups of pupils.

Tanner (1961) said that there appeared to be a greater loss in knowledge of subject matter in groups of students taught by television when compared with conventionally taught students.

Cropper and Lumsdaine (1961) modified the method of presentation of a television lesson to allow pupils to make active responses during the lesson. They called this a programmed lesson. The results of the investigation showed the pupils taught by the programmed television lesson, to have scored higher than those taught by conventional televised instruction both on an immediate achievement test and on a later test of retention of learning. The investigators commented that: ‘ This lends some support to the value of active response for learning and retention ’.

Kumata (1960) claimed that no significant differences existed between the retention of learning of television taught pupils and that of conventionally taught pupils. However, Schramm (1962), said that there were no solid conclusions to say that television helps or hinders retention.

Barrington (1964), some four months after the end of the teaching period in his investigation, found no actual difference between the loss of learning of pupils taught by television plus laboratory work and those taught by conventional laboratory lessons. Both groups appeared to have retained all their learning. No significant difference was found between the loss of learning of pupils taught by television in the classroom and those taught by conventional classroom lessons. Both groups suffered a similar loss of learning—about 30%. These findings appear to support the opinion of Cropper and Lumsdaine concerning ‘ active response ’.

5. Research into pupils’ attitude to television teaching

Heminghaus (1957), described the use of a questionnaire in St. Louis schools to obtain information about the attitude of pupils to being taught by television. The responses to the questionnaire indicated that the pupils believed that they would learn more from personal instruction although there was no significant evidence to support this view.

Bobren (1958 and 1960) concluded that television students had a negative attitude to learning from television.

W. H. Allen (1960) reported that although television students may have an unfavourable attitude to learning by television their performance was as good as that of conventionally-taught students.

Kumata (1960) noted that elementary pupils were enthusiastic about television lessons, high school pupils were less so, and university students were unenthusiastic. He also found that pupils who had been taught by television had a more favourable attitude to television lessons than those who had not had any experience of television teaching. He suggested that pupils become conditioned to what is the standard educational procedure.

Brish (1960), reporting on a survey carried out in Hagerstown, said that pupils had a very favourable attitude to television lessons.

Barrington (1961) said that the results from a simple unscaled questionnaire completed by pupils who had followed television science lessons indicated that the pupils had a negative attitude to the television lessons which were of the enrichment type but that they were asking for direct teaching by television. Their achievement was similar to that of conventionally-taught pupils.

Tanner (1961) and later Meaney (1962) said there appeared to be little direct relationship between the students' attitude towards television instruction and their achievement on the courses. Meaney also found, as Kumata had noted earlier, that pupils become more favourable to television teaching after following a series of television lessons.

Schramm (1962), like Kumata, found that elementary school pupils were enthusiastic about television lessons, high school pupils were less so, and college students were equivocal and sometimes unfavourable. He noted that there were indications that the particular subject and the television teacher determined the pupils' attitude.

Jacobs and Grate (1962) included an attitude study in their research. They found that the pupils were favourable to television teaching. They noted that above average pupils appeared to be less favourable to television instruction then pupils of below average ability. Westley and Jacobsen (1963) concluded that pupils had a favourable attitude to television. Their most striking finding was that the television taught pupils tended to rate their own teacher more highly than did the non-television pupils and more highly than they rated the television teachers.

Daines (1962) also found that pupils taught by television showed a favourable attitude to being taught by television. These findings supported conclusions from his earlier investigations.

Barrington (1964) found that pupils taught by television both in a classroom and in a laboratory had favourable attitudes to television lessons and that there was no significant difference between the attitude of the groups. He also reported that the more able the pupils, the less favourable their attitude to television lessons although the attitude of the pupils did not appear to affect their performance on achievement tests.

Janes (1964) attempted to identify student traits which might encourage positive or negative response to televised lectures. A measure of authoritarianism, an intelligence score, and a measure of self-confidence were used as measures of traits suitable for the investigation. The results indicated that all traits had positive correlations with preference for television lectures. Authoritarianism correlated slightly, intelligence correlated more highly and self-confidence correlated most highly with a favourable attitude to televised instruction.

6. Summary and findings from instructional television researches

From this survey of researches in the field of instructional television, the following conclusions appear to emerge:

1. There appears to be no significant difference in achievement between pupils taught by television and by conventional methods.
2. The few experiments aimed at measuring the intangibles of the learning situation have also failed to discover significant differences between television and conventionally taught pupils.
3. Follow-up work after a television lesson is vital.
4. Television lessons hold the pupil's attention.
5. The evidence indicates that the optimum length of a television lesson is about 25 minutes.
6. Sound should be closely related to what can be perceived.
7. The effect of television teaching on the achievement of pupils of different levels of ability has not been resolved.
8. Pupils favour two-way communication between the classroom and studio. But the absence of two-way communication does not affect the level of achievement of the pupils. There is evidence to indicate that it may impede the learning taking place.
9. It has not been satisfactorily resolved whether television teaching helps or hinders retention of learning. When pupils are engaged in making active responses during a television lesson, retention is much improved.
10. The attitude of pupils to television instruction is affected by the subject being taught and by the television teacher.
11. The less able pupils and younger pupils are enthusiastic about television lessons. The pupil's attitude becomes unfavourable as he climbs the educational ladder from primary to secondary and higher education. However the attitude of the pupil towards television lessons does not appear to be related to his level of achievement in the subject being taught.
12. Pupils appear to favour the direct teaching type of lesson rather than the enrichment type of lesson.

There is still a great need for research to be carried out aimed at improving the presentation of television lessons and at discovering the most effective way of using television for teaching. Objective evidence of the effectiveness of television lessons would be welcomed by teachers eager to employ new methods of teaching and should also be appreciated by those not easily persuaded to change from established methods of teaching.

References

A.B.C. (1963). *Under Observation.* A.B.C. T.V.

A.V.C.R. (1961). 'Effect of Eye Contact on Television Instruction', Abstract in: *A.V.C.R.. No.* 2.

Allen, W. H. (1960). 'Audio Visual Communication'. *Ency Ed. Research.*

Amaldi, A. (1960). *Television for School Science.* O.E.E.C.

Ambrosino, M. J. (1957). 'The Schenectedy Experiment', Abstract in: *Audio Visual Communication Review, No.* 4.

Anderson, G. R., and Vandermeer, A. W. (1955). 'Slide Rule Instruction by Television', Abstract in: *A.V.C.R. No.* 2.

Associated Rediffusion (1962). 'The First Four Years'. A.R. T.V.

Barrington, H. (1961). 'A Comparative Study of The Achievement of Pupils who have studied Physical Science Topics by Television or Laboratory Lessons', *Dip. Ed. Psych.* University of Manchester.

Barrington, H. (1964). 'The teaching of elementary science by television and other methods', *M.Ed.* Univ. of Manchester.

Baxter, J. (1960). 'Continental Classroom', T.V. for Science Teaching. *O.E.E.C.*

Becker, S. L. (1964). 'Interest Tension and Retention', *A.V.C.R.* Fall.

Bobren, H. M. (1960). 'Students Attitudes to Instructional Television', *A.V.C.R., No.* 6.

Boeck, C. H. and Washton, N. S. (1961). 'Science in Secondary School', *Review of Educational Research,* June.

Bohlen, A. (1960). 'History and Theory of Mass Media in Germany', *Year Book of Education.*

Bridges, C. C. (1960). 'An Attention Scale for Evaluating Educational Television Programmes', *Journal of Ed. Research,* Dec.

Brish, W. M. (1960). 'C.C. T.V. for Science', *Year Book of Education.*

Brish, W. M. (1961). 'C.C. T.V. Project Washington County'—Fact Sheet. Board of Education, Washington County, Hagerstown, Maryland, U.S.A.

Brish, W. M. (1964). 'Washington County C.C. T.V.'—Report. Board of Education, Washington County.

Brown, C. (1954). 'Teaching by Television', *Schoolmaster,* 18. 6.

Bull, C. R. (1960). 'Mass Media in Australia', *Year Book of Education.*

Carpenter, C. R. (1960). 'Television'. *New Media in Education.* Edling, J. V., Ed.

Carpenter, C. R., and Greenhill, L. P. (1958). 'Investigations of C.C. T.V. for Teaching University Courses', Report No. 2., Pennsylvania State Univ.

Cassirer, H. R. (1960). *Television Teaching Today.* UNESCO.

Central Office of Information (1963). *Learning by Television.* R.S.N. 5517, C.O.I.

Cincinnati Board of Education (1958). Report in: *Review of Ed. Research, No.* 1.

Daines, J. W. (1962). 'Science Without a Specialist Teacher', *Bulletin of Institute of Education.* Univ. Nottingham, Sept.

Daines, J. W. and Neilson, J. B. (1963). 'Objectivity and Subjectivity', *Bulletin of Institute of Education.* Univ. of Nottingham, October.

Daniels, J. C. (1959). 'Experiment in Television for School', *Bulletin of Institute of Education.* Univ. of Nottingham, Jan.

Deutschmann, P. J., Barrow, L. C. and McMillan, A. (1962). 'Efficiency of Different Methods of Communication', *A.V.C.R., No.* 3.

Dieuzeide, H. (1960). *Teaching Science Through Television.* O.E.E.C.

Divitt, J. J. (1963). Intellectual Stimulation of Gifted Pupils Through T.V. Instruction', *A.V.C.R., No.* 5.

Dowell, E. C. (1956). 'Experiment in Teaching by T.V.', *A.V.C.R., No.* 3.

Dreher and Beatty (1960). Report on Ed. T.V. in: *Review of Ed. Research, No.* 3.

Erickson, C. G. (1960). 'Closed and Open Circuit for T.V. Courses', *Year Book of Education.*

Ford Foundation (1961). *Teaching by Television.* Ford Foundation and Fund for Advancement of Education.

Glasgow Education Committee. 'Glasgow Schools T.V.', *Times Ed. Supplement,* 22.11.63.

Greenhill, L. P. (1959). *T.V. in University Science Instruction.* Pennsylvania State University.

Gropper, C. L. and Lumsdaine, A. A. (1961). 'An Experimental Comparison of a Programmed T.V. Lesson with a Conventional T.V. Lesson', *A.V.C.R., No.* 6.

Guillin, M. (1960). 'Science Programmes for the 12/15 years Age Group', *Television for School Science.* O.E.E.C.

Harris, C. O. (1964). 'Learning of Relevant and Irrelevant Information', *A.V.C.R.,* Spring.

Hatano, K. (1960). 'Mass Media and Education in Japan', *Year Book of Education.*

Hayman, J. L. (1963). 'Viewer Location and Learning in Instructional T.V.', *A.V.C.R., No.* 3.

Hayman, J. L. and Johnson, J. T. (1963). 'Exact versus Varied Repetition in Educational T.V.', *A.V.C.R., No.* 4.

Heminghaus, E. G. (1957). 'St. Louis Experiment', *School Review, No.* 2.

Holmes, P. D. (1962). 'On Understanding T.V. Significant Differences', *A.V.C.R., No.* 4.

Jacobs, J. N. and Bollenbacher, J. (1960). 'Retention of Subject Matter in T.V. Biology', *A.V.C.R., No.* 6.

Jacobs, J. N. and Grate, J. (1962). 'Teaching Sixth Grade Science by T.V.', *Elementary School Journal.* Univ. of Chicago, November.

Janes, R. W. (1964). 'Pre existing Attitudes of College Students to Instructional T.V.', *A.V.C.R.,* Fall.

Kanner, J. H. (1957). 'Future Trends in T.V. Teaching and Research', *A.V.C.R., No.* 4.

Kanner, J. H. and Marshall, W. P. (1963). 'The Improvement of Training by Television', *A.V.C.R., No.* 5.

Kanner, J. H., Runyon, R. P. and Desiderato, O. (1955). 'Television as a Training and Educational Medium', *A.V.C.R., No.* 3.

Kumata, H. (1960). 'A decade of Teaching by T.V.', *Impact of Educational Television.* University of Illinois.

Love, E. (1963). 'Competition the Spur', Teacher T.V. Supplement, *The Times.*

Maeda, Y. (1962). 'T.V. for Teaching', British Assoc. Granada Lecture. Granada T.V.

Matthews, M. (1963). 'Growing Interest and Experiment', *Times Ed. Supplement*, 22.2.

Meaney, J. W. (1962). *Televised College Courses.* Ford Foundation, Oct.

Mercer, A. and Becker, S. (1955). 'The Disenchantments of Educational T.V.', *A.V.C.R., No.* 3.

Miami Report (1960). *Review of Educational Research, No.* 4, p. 354.

Ogren, G. (1960). 'Mass Media in Swedish Education', *Year Book of Education.*

Pentalakis, S. N., Bower, B. D. and Jones, D. H. (1962). 'Convulsions and Television Viewing', *British Medical Journal*, 8th September.

Perry, A. (1960). 'Teaching by T.V. in Today's Schools', *Educational Forum*, May.

Postman, L. (1961). 'Human Learning and Audio Visual Education', *A.V.C.R., No.* 5.

Puglisi, M. (1962). 'Telescuola', *Bulletin of Institute of Education.* Univ. of Nottingham, Sept.

Purdue Report No. 9 (1960). *Review of Educational Research, No.* 4.

Rehage, K. J. (1959). 'How Fares Educational T.V.?', *Elementary School Journal*, December.

Roy, W. (1959). 'A Year of School T.V.', *Bulletin of Institute of Education.* Univ. of Cambridge, June.

Royds, A. (1963). *Schools T.V. and the teaching of elementary science.* Ph.D., University of Manchester.

Scanlon, J. J. (1960). 'C.C. T.V. Possibilities and Limitations', *T.V. for School Science.* O.E.E.C.

School Broadcasting Council (1959). B.B.C. School Television Broadcasting.

Schramm, W. (1962). 'Learning from Instructional Television', *Educational Television—The Next Ten Years.* Stanford Univ.

Seibert, A. (1958). 'Experiments at Purdue University', *Review of Educational Research, No.* 2.

Seldes, G. (1957). 'Television and Education', *Confluence*, Summer.

Shimberg, B. (1957). 'Effectiveness of T.V. in Teaching Home Nursing', Abstract in: *A.V.C.R., No.* 2.

Siepmann, C. A. (1952). *Television and Education in U.S.A.* UNESCO.

Signal School, U.S.A. (1957). 'Teaching by Television', Abstract in: *A.V.C.R., No.* 2.

Skornia, H. J. (1960). 'Radio and T.V. in Education in U.S.A.', *Year Book of Education.*

Suchodolski, A. (1960). 'Educational Tradition and Mass Media in Poland', *Year Book of Education.*

Sunesen, B. (1960). 'Recent Developments in Denmark', *Year Book of Education.*

Tannenbaum, R. (1960). 'Instruction through Television', *Encyclopaedia of Educational Research.*

Tanner, D. (1961). 'Needed Research in Instructional T.V.', *School Review*, Autumn.

Travers, B. H. (1960). 'Some Thoughts on Educational T.V.', *Australian Journal of Education*, Vol. 4, No. 3.

Trenamen, J. (1959). *An Investigation by Statistical Methods of the Effective Communication of Educative Material, etc.* Ph.D., Bodleian, Oxford.

U.S. Navy (1957). 'Instructional T.V. Research Reports', Abstract in: *A.V.C.R., No.* 2.

Vernon, M. D. (1953). 'Perception and Understanding of Instructional Television Programmes', *Brit. Journal of Psychology*, 44116.

Weiss, J. K. (1957). 'Hagerstown Experiment', *School Review, No.* 2.

Westly, B. H. and Jacobsen, H. K. (1963). 'Instructional Television and Pupils' Attitudes', *A.V.C.R., No.* 3.

Williams, C. M. (1962). 'Re-examination of "No Sig. Diffs." that Instructional Television Studies Report', *A.V.C.R., No.* 4.

Williams, T. Reports in: *Exploration*, August 1954, and in *Review of Educational Research*, 1960, No. 4.

Wilson, J. (1962). 'Educational Television and Africa', *Teacher Education*, May.

Wolgamuth, D. (1962). 'A Comparative Study of Three Techniques of Student Feedback in T.V. Teaching', Title VII Abstract in: *A.V.C.R., No.* 3.

SOME PROBLEMS INVOLVED IN THE EXPERIMENTAL COMPARISON OF TEACHING METHODS

by J. D. Williams

Introduction

IF flawless educational research and efficient communication of its results were usual, then perhaps little harm would come of accepting at face value the conclusions drawn by those who engage in it. Usually, however, a close examination of the precise conditions under which an investigation has been carried out will reveal that such conclusions should be heavily qualified before they are allowed to influence educational practice.

Much of that research which would appear to be the most pertinent to the teacher's problems has been carried out under 'live' teaching conditions—conditions that are as representative as possible of the usual classroom situation. While this kind of research potentially has the merit of doing justice to the full complexity of the teaching situation, it presents the investigator with some awkward problems. In one way or another, many of these problems derive from the fact that, before he can come to grips with those relations between variables in which he is interested—and these may be decidedly slippery—the investigator must find his way through the plethora of insistent and interfering influences that bear upon—and sometimes bear *away*—these relations. Some unwanted influences can be removed, some can be neutralised, and for some it is possible to make allowances, but, unfortunately, it is not unusual for certain of such influences to resist all efforts either to dismiss or to assess them.

Usually, in his research report, the investigator will acknowledge the effects of these influences; he may point out departures from a simple basic research design, actual inadequacies in design, respects in which the results of the investigation must be qualified or doubted, and so on. However, where such reports are available, the detailed reservations they contain may well be overlooked, for, typically, the practising educationist has little interest in the precise procedure adopted in an investigation, and will tend to devote the greater part of his attention to that section of a report which has the most immediate relevance to teaching-practice: the summary and conclusions. But frequently the state of affairs is sorrier than this, for, typically, such reports are *not* available, and the educationist will be quite without convenient access to an account of the shortcomings of design and execution that should determine his interpretation of the findings of an investigation. Most unfortunate, and most usual, is the situation in which the results reach the educationist only after they have passed through channels of communication which have filtered off all meaning but that which can be contained in the very simplest of statements. A statement like 'X causes Y' is easy to communicate, to interpret, and to adopt as a guide for teaching-practice, but it hardly constitutes a valid contraction of some such original as 'Under conditions A, B and C, but not D, E and F, we obtained a slight correlation—which, however, was not significant at the .05 level—between X and Y'.

The assumption underlying the present article is that it is sometimes possible, when one knows the general nature of the investigation in question, to read between the lines of over-abbreviated accounts of research findings, and, in interpreting these findings, to bear in mind certain kinds of difficulty that can be expected to dog certain kinds of research enterprise. In particular, the purpose of the article is to describe some of the difficulties that may be expected to arise in the design and execution of an experimental comparison of teaching methods, and to indicate how, and to what extent, it is likely that an investigator will be able to cope with such difficulties. By this means we hope to incline the educationist, and even to a degree to *enable* him, to consider more critically statements of the findings of investigations of both this kind and others. While it is hoped that the article will be of general relevance, the points contained in it will be made with reference to a particular experiment* which has recently been performed by the N.F.E.R., and about which, consequently, we are in a position to provide, from behind the scenes, detailed information about the obstacles that have threatened to obstruct its course, about the measures taken to remove or circumnavigate these obstacles and about the respects in which such vicissitudes as these are likely to affect its outcome.

The main object of the experiment was to compare three 'structural' methods of teaching arithmetic—the Cuisenaire, the Dienes and the Stern—with traditional methods. All three experimental methods were used in the first and second years of the junior school, and, to a smaller extent, the Cuisenaire and Stern methods were used in the infants' school. Two experimental designs were employed—an 'external-control' and an 'internal-control' design. In the case of the 'external-control' design, the experimental children (that is, those using structural methods) were drawn from different schools from those of the control children (those children using traditional methods). In the case of the 'internal-control' design, groups of both kinds of children came from each school, but the experimental children from a later generation than that of the control children. As instruments of evaluation, tests of mechanical, problem, and 'concept' arithmetic, together with attitude questionnaires, were administered at the beginning and end of each experimental period. Other relevant information, concerning the general ability of the children, teaching circumstances, etc., was obtained by means of teacher-questionnaires, and, in some cases, by means of school visits. Further particulars of the experiment will become apparent in the following discussion of the problems it posed.

At first sight, an investigation of the above kind may seem to be straightforward enough. We take two groups of children, teach one by the experimental method, and one by the control method, then compare their performance. On closer inspection, however, the situation can be seen to bristle with problems of execution, of experimental engineering and even of definition. We have been able to cope quite adequately with some of the problems, and include an account of them here merely in order to indicate that such problems exist; but there are certain problems (some of which researchers seem to have resigned themselves to 'living with', and which they habitually omit to mention in their reports) to which it seems to be impossible, under the circumstances of operation of this kind of research, to find a fully legitimate solution. We shall devote a fair amount of attention to the analysis of such problems, for these set limits both to the scope of investigations of this kind, and to the relevance of their results.

---

*Although this experiment is now very near to completion, results are not yet available. However, a brief description of its main features may be obtained from the N.F.E.R. upon application.

Teaching methods and impure experimental entities

The three methods of teaching arithmetic, which, in this investigation, we have chosen to assess—the Dienes, the Cuisenaire and the Stern—are standard methods. One of the factors that determined our selection of these methods was the availability, in the case of each, of a well-developed scheme of work. Now, in relation to the instruments with which we shall be assessing the methods, we could say of such schemes of work that they are 'experimentally complex'. That is, each scheme embodies a variety of features which may be responsible for the effects we discover, and, in some cases, features which may mask the effects of others. As an illustration of this point, let us take the Dienes method. This method has several pedagogically-interesting features: concrete apparatus is used, ideas must be abstracted from a variety of number-systems, pupils are encouraged to discover relations for themselves. Since all these features operate at once, and each to an unascertainable degree, how can we establish what part each plays in the production of the results that derive from using the Dienes method? Perhaps the method would produce its effects without the use of concrete apparatus, or perhaps this apparatus is their sole cause. Perhaps the use of a variety of number-systems is, to the duller children, as confusing as the use of concrete apparatus is elucidating, so that the beneficial effects of using apparatus are masked.

As a source of information that might guide us in improving our mathematics teaching, a study of this kind—one in which we are assessing complete teaching methods, each consisting of a complex of relevant components—is far from maximally efficient. If we could vary the relevant components, in order to establish the contribution of each to the final effect, then we might gain information which would enable us to improve our teaching methods. As it is, we shall be able to judge each method only in terms of its overall effect. We may, as a result, be able to say how, in certain ways, a particular method compares with others, but we shall be denied specific knowledge of those aspects of the method which are responsible for its merits and deficiencies.

However, it is a consolation that the results of this kind of investigation will be directly relevant to the practices of many teachers, for, while the Cuisenaire, Stern, and Dienes methods do not make ideal *experimental* entities, they certainly constitute entities for the practising teacher, and any conclusions concerning the overall effectiveness of each method will provide him with valuable assistance in deciding between methods.

Equating schools and pupils

Naturally, a comparison between experimental and control children will be legitimate only if it is possible to discount those differences between groups that are due to such factors as pupil ability and school setting. In designing the experiment we have used two main devices for minimising or controlling the effect of such factors, but neither device can wholly rule out the possibility that these factors will affect the results.

The first device consists in running both experimental and control groups in the same school. The obvious way of doing this—to compare parallel streams within a school—was difficult to arrange, for few schools would be willing to disrupt their organisation to an extent that would permit such an arrangement. Even had it been possible to arrange this, there would have arisen a further difficulty: the experimental methods have rather distinctive teaching policies, and these may have tended to filter into the control classes, thus reducing the contrast between methods.

In part of our experiment we did this in a less direct form. We compared a generation of children who had used control methods with a subsequent generation who had used experimental methods. This form was perfectly congenial to the schools, for it required no class reorganisation. There was no risk of contamination by the experimental method, for this method was introduced after the controls had been tested. In this way, we were able to compare two groups of children who had been drawn from the same neighbourhood and taught in the same school setting. However, this resort was not without its snags, the most important of which was that different generations of children developed at different points in a climate of mathematics-teaching that was rapidly changing. At the time of the experiment there was a general awakening of interest in primary mathematics-teaching, and it is quite possible that, even within the span of a year, the effect on teaching standards of this awakening was considerable. If so, we could hardly claim that control and experimental pupils were subjected to comparable learning conditions. Fortunately, we shall be able to establish whether our results have been influenced by this factor, and, if so, to what extent; for in another part of the experiment, we have tested pupils who come from the same generation as those in the experimental group, but who have not used an experimental method.

The second device was to equate experimental and control groups by testing them before the experimental period began. With the groups equated in this way, one could perhaps assume that differences found at the *end* of the experimental period would be due to the method of teaching used. However, during the experimental period it could well be that influences other than teaching-method operated in such a way as to introduce differences between the groups. Apart from teaching-quality, which will be considered in the following section, such factors as school facilities, parental encouragement and differential maturation could have introduced an imbalance. While we have been unable to eliminate the effect of such factors, we have done our best to record the extent of the most relevant of them, and shall be able to take these into account in the analysis of the results of the experiment.

Teacher-ability

Among the factors that may affect the pupil's attainment in arithmetic, teacher-ability is perhaps the most important. Where the experimental and control groups are taught by different teachers, effects due to this factor could well mask those due to the teaching method used. One might hope to eliminate the effect of teacher-difference by using the same teacher for both experimental and control groups. We have managed to establish this condition in some of those schools in which we have compared different, but comparable, generations. But even under this condition we cannot be sure that the groups were matched for teaching-effectiveness, for it is clear that the effectiveness of a teacher may depend upon the method he uses. For example, a teacher who is authoritarian may be very good at teaching by a very formal, traditional, method, but may be quite unable to cope with the permissive kind of teaching situation that is supposed to obtain where the Dienes method is used.

One reason why teacher-ability is likely to be so important a factor, is that each teacher affects the performance of up to 40 children, so that, although we may have many children in our experimental and control groups, thereby increasing the likelihood that these groups will be balanced with regard to child-ability, we shall still have relatively few teachers in our groups, which are therefore much less likely to be adequately balanced with regard to teacher-ability. In order to increase the number of teachers in the groups without increasing beyond the limits of our budget the number of pupils to be tested, we have, in part of the experiment, tested only every *other* child in the participating classes. If there had been more schools available for purposes of experimentation, we might have employed this method of reducing our sample more consistently, thus reducing even further the large child-teacher ratio.

Teacher-selection

Ideally, in selecting an experimental sample, it should be possible to contact large numbers of teachers who would be willing to co-operate in the experiment in whichever capacity the experimenter chose. The experimenter would then assign teachers to experimental and control conditions in accordance with standard sample-selection procedures. However, in the case of the present experiment, the budget was so small that the experimental schools were required to buy their own apparatus for use with the experimental methods. This source of discouragement reduced so drastically the number of schools that was available for inclusion in the experimental group, that we had to include in this group nearly all of those schools that were willing to meet the expense involved. Again, it was difficult to find schools that would have been willing to pay for expensive apparatus but not disinclined to join the control group, for the willingness to meet such an expense required a level of enthusiasm for the particular experimental method concerned that was unlikely to be equalled by the school's enthusiasm for research. Hence the experimental group is composed largely of children from schools that have been willing to make a considerable financial sacrifice, while—in the 'external control' section of the experiment—the control group is composed of children who do not necessarily come from such schools.

Obviously, this is a condition that is likely to have favoured the selection of particularly enthusiastic teachers for the experimental group. Again—at least in the case of the Dienes experimental group—the teachers must have had a fairly sophisticated knowledge of elementary mathematics in order to have been able to appreciate the merits claimed for the experimental method. Both these biases of selection are likely to have favoured the experimental over the control group. It should be pointed out, however, that in that part of the experiment in which this bias could have operated, the control teachers, as well as the experimental teachers, are likely to have been both enthusiastic over mathematics, and relatively enlightened, for all teachers were drawn from a population who had been voluntarily attending a series of meetings on the teaching of mathematics.

Even when the experiment was under way, selection that was related to teacher-quality is likely to have operated in two ways. First, many teachers eventually reduced the amount of time they spent on the use of the experimental method. Perhaps such teachers tended to be less conscientious or talented than those who persisted in the full use of the method. If this were so, the results of the experiment may show a correlation between degree of use of the experimental method, and pupil-attainment; it may be erroneous to conclude from such a result that, in itself, a greater degree of use of the method produces higher attainment. Second, teacher-characteristics may well be related to the teacher's continued use of the experimental method throughout the full three years of the experiment. Since some teachers dropped the method altogether, the experimental group may have become progressively more saturated with teachers whose enthusiasm for and compatibility with the method was great, and whose teaching ability was high. This tendency could affect the results in two ways: high attainment may be the result of teacher-selection rather than of the method being used, and apparent teacher-improvement over the period of the experiment may also be an artefact produced by teacher-selection.

While most of the factors that can be seen to have operated in teacher-selection are likely to have favoured the experimental group, there is one which may well have placed this group at a disadvantage. In the case of some experimental schools, participation in the experiment and adoption of the experimental method was obviously the result of the corporate decision of both headmaster and staff, but in others it was by no means always obvious that the assistant teachers, who, during the experiment, actually *operated* the method, had entirely wished to adopt it or were entirely pleased with its use. It may be, of course, that such conditions obtained in the control schools, also, but perhaps it is legitimate to suspect that a situation in which an unusual and new method has just been introduced into a school is one that might conduce to a heightened degree of dissatisfaction and bewilderment among teachers who have had no say in the adoption of the method.

Teacher-experience

Where a new teaching method is being compared with one that is more established, the new method is liable to stand at a disadvantage in two important respects. First, it may not have reached its 'mature' form, and in its infancy may be in a relatively crude state. Of our three experimental methods, it is safe to say that the Cuisenaire and Stern were 'full-grown' at the time of the experiment, but the Dienes method, partly because it involved more revolutionary innovations than the others, and partly because it had had less time in which to develop, was still having many of its teething troubles during the course of the experiment.

Second, the teacher using the new method is likely to be a novice in its use. It would not have been possible, at the time at which the experiment was conducted, to obtain teachers who were experienced in using the Dienes method, for this method had been in existence for far too short a time. While the Cuisenaire and Stern methods had been in a fairly fully-developed form for a considerable period, their use was not sufficiently widespread, at the time at which the experiment was launched, for it to be possible to find many schools both using these methods *and* in a location that was convenient for the running of an experiment.

Quite apart from this difficulty in finding experienced teachers, the design of part of the experiment made it virtually impossible for such teachers to participate. Where we compared successive generations in the same school, teachers who had previously been teaching conventionally were required to switch over to the new method, so that earlier and later generations of children, from the same school and under the same teacher, could be compared after having been subjected to different teaching methods. Usually, this meant that the teacher had the opportunity to use the experimental method for only one or two or three years before its effectiveness was compared with that of methods that the same teacher had been using for perhaps the whole length of his previous career.

Is there any special reason why experience should be an important factor in the operation of the experimental method? Unfortunately, it appears that experience is likely to affect considerably both the efficiency and the accuracy with which these methods are employed. All three methods, but particularly the Dienes, require of the teacher a capacity for coping with a kind of classroom organisation with which, if he has previously taught conventionally, he is unlikely to be very familiar: children must be allowed for some of the time to work in groups; they will tend, since they work at their own pace, to be, for most of the time, at different points in the syllabus; the use of apparatus can introduce into the classroom a great deal of disorganisation; the teacher must learn the subtle arts of coaxing the pupil to discover, instead of merely presenting him with information.

What are the consequences of teacher-inexperience? There must be many such consequences, but some have been more evident than others. First, the operation of an unfamiliar method may place upon the teacher a degree of stress to which he is not accustomed. This is likely to encourage him to regress to familiar, and therefore easier, techniques of teaching—which he may do either overtly or in ways that are more difficult for the experimenter to detect. Second—partly because the teacher will tend to retain teaching-objectives that are more appropriate to his previous method, and partly because, even in relation to more suitable objectives, the standard of his teaching will not, while he is accommodating himself to a new method, be at its highest—he is unlikely to find much encouragement from the results of his efforts. This, also, will incline him to return to his former techniques. Third, until the teacher has been through a sequence of lessons at least once, he cannot hope to appreciate the relative emphasis to place upon each lesson, or to know how to use earlier lessons as a preparation for those that are to come later. Fourth, almost every teacher finds great difficulty in estimating how long a course of lessons will take. Because of this last difficulty, many teachers fail to complete the work that is meant to be covered in a given period, and there is a consequent danger that the pupil's eventual test-performance will be affected. Happily, this difficulty was foreseen, and, while the design of most of our tests is such that lack of syllabus-coverage will affect only minimally the pupil's performance, there has been included one test which should reveal with particular clarity the extent to which this determinant of performance has operated.

One of our means of ascertaining the effect of teacher-inexperience is to compare the performance of children who have been taught by a teacher who has had one year's experience in using the method, with that of a comparable group of children who have been taught by the same teacher after he has used the method for two years. Where there is significant improvement, there must have been room for it.

Teacher-constancy

Where a teacher in the experimental group drops out, it is unlikely that he will be replaced by one who is acquainted with the relevant experimental method, for these methods are fairly uncommon. Therefore, in the experimental group, it is likely that where a teacher leaves the experiment, the class will either leave with him, or be inadequately taught by a replacement who is a novice to the experimental method. The control group will probably suffer less in this respect, for it is more likely, in the case of a conventional teaching-method, that a replacement teacher will have the kind of experience that will enable him to take up the threads of his predecessor's teaching.

This lack of teacher-constancy will be more pronounced in schools where the teachers are young, and therefore mobile. Unfortunately, there are many such schools in our sample—perhaps because young teachers are more interested in new methods, and because the kind of head teacher who is interested in research and new methods is also likely to welcome youthful staff.

The experimental group's need for teacher-constancy is not *entirely* to its disadvantage, for, where this need is recognised, circumstances are often arranged so that it will be satisfied. Headmasters in experimental schools often make arrangements for an experimental class to be taught by the same teacher in both first and second years, for often there is only one teacher available who has the necessary skills to continue the tuition of this class by the experimental method. This tendency is likely to favour the experimental group over the control group.

It may be thought that two opposing influences such as those mentioned above can be assumed to nullify one another. If their effects could be measured, and were found to be equal, we could, indeed, discount them. But as it is, we have no way of comparing the magnitude of these effects, so must acknowledge that our belief in some aspects of the validity of our experiment must hinge upon our readiness to assume that these effects are negligible.

Class-constancy

The above problems of teacher-constancy are closely paralleled by problems of what we could describe as 'class-constancy'. Where new children join a control class, they are unlikely to find too much difficulty in adjusting to the new learning-conditions, for these conditions will probably bear a close resemblance to those of their previous class. By contrast, the problem of adjusting to the peculiar learning-conditions of one of the *experimental* classes is likely to be quite serious, for the odds are that new children will not have had previous experience with so uncommon a teaching-method. This difference can be expected to place the experimental group at a disadvantage.

Where the headmaster sees a threat to the continuity of his pupils' learning, he is likely, as in the case of *teacher*-constancy, to take protective measures. In this case, he will try to ensure that the experimental class remains intact. For the experimental class, this will be a mixed blessing, for while it will make for stability of class-organisation, and will help to ensure that the standard of the class is not reduced by the inclusion of pupils whose learning-career has been disrupted, it will, at the same time, preclude the re-sorting of pupils that is often necessary for the improvement of class-organisation.

The specification of control conditions

A main problem in conceiving the experiment was that of deciding with what to compare the experimental group. Occasionally, in the above comments, we have referred to the methods used in the control groups as 'conventional' or 'traditional'. These are extremely ambiguous terms, and, probably, if we managed to give them a precise meaning, we should find that they applied to very few indeed of the methods used in our control classes. After a certain amount of deliberation, we decided that the methods used by the classes of our control group should be as representative as possible of general teaching practice. Hence, we have debarred from this group no method excepting structural methods of the kind that we have included in our experimental group.

The main issue involved in this decision is that of whether we should sacrifice precision for relevance. If we had provided the control schools with a prescribed syllabus, or had selected them according to a rigorous prescription, then we should have been in a position to say exactly with what the experimental methods were being compared. Then, perhaps, we could have come to some very definite conclusions about the components of these methods. For example, if we had prescribed that control methods should entail the use of a variety of experiences from which concepts could be abstracted (multi-embodiment), then we could have attributed differences between the effects of control and experimental methods to features of the methods other than that of multi-embodiment. On the other hand, there would have been no way of ensuring that our prescription for teaching was representative of the kind of classroom situation to which most practising teachers would eventually wish to relate the results of the experiment. A further important factor is that prescription is *in itself* likely to have a peculiar effect upon the teaching situation, for different teachers are likely to find different methods the most effective, and the prescription of one particular method may well severely limit many teachers.

As it is, we shall be able to specify to *some* degree the characteristics of our control classes, for we have gone to a great deal of trouble to record and classify the teaching techniques used in them. Thus, while we have a representative group of control classes—which we shall be able to use for making overall comparisons that will be relevant to classes in general—we shall at the same time be able to make more detailed comparisons which will give us a more precise insight into those aspects of both experimental and control methods that are responsible for the results we obtain.

The specification of experimental conditions

Although it cannot be said that we are unsure of the essential nature of any of the methods to be used in the experimental group, there may still be a great deal of variation in the degree and kind of use to which each of these methods is put. We are faced, therefore, with a problem similar to that with which we had to deal in the case of the control group—that of deciding how far we should prescribe the manner and degree of use of the experimental methods. Again, we must consider this problem in terms of the degree to which we are willing to sacrifice experimental precision in favour of the relevance of our results. If we were to set strict limits to the teacher's choice of how, and to what extent, to use an experimental method, then we could expect to know precisely to which conditions our results would relate—but we should have no right to claim that these conditions represented the usual conditions of operation of the method. We should also lay ourselves open to the charge that we were impairing the teacher's efficiency by cramping his style, and in this way, also, experimenting in a situation that was not representative of usual teaching-practice.

We decided to sacrifice precision of specification of the conditions in which the methods were to be used, so that teachers would be able (as they normally would, outside of the experiment) to integrate the experimental methods with their normal techniques of teaching, in a way that they considered to be most effective. This does not mean that we are wholly ignorant of the manner and degree of use of the experimental methods: questionnaires and interviews have provided us with detailed information concerning the manner in which teachers chose to use these methods. It does mean, however, that we have been unable to plan the experiment in such a way that we could be sure of having certain numbers of children taught in certain pre-arranged ways. It means, also, that we have had a greater drop-out from the experimental group than we otherwise might have had, for some teachers found that they could teach most effectively when using an experimental method to a negligible degree!

Unplanned changes in experimental conditions

It is impossible to anticipate in detail, at the outset of a three-year experiment, precisely what will happen during its course—naturally, the longer an investigation takes, the more difficult it will be, at the planning stage, to take into account the changes in circumstance with which it may be confronted. This is particularly so in the case of an experiment such as this, which has been carried out in a 'truly representative' teaching context. While interested in the experiment, and, in nearly every case, as co-operative over it as could have been hoped, the participating teachers naturally considered that their first duty was to their pupils, and, where good experimentation came into conflict with good teaching, the latter was usually permitted to prevail. Had the experiment been carried out in special experimental schools, or even had our budget been generous enough to allow us to save schools the expense of purchasing for themselves the necessary structured apparatus, then, maybe, we could have expected to have been able to control the course of the experiment more closely.

The most obvious, and in some ways, most saddening kind of change, is due to the depletion of the number of children in the sample. The commonest cause of this is that the teacher who has experience with an experimental method will leave the school, and his replacement will be regarded as having insufficient knowledge of the method to continue its use. Sometimes, a change in headmaster will precipitate change from an experimental method. Since the experimental methods demand uncramped teaching conditions, such factors as temporary work on the alteration of the school building can lead to a cessation of the use of these methods. Sometimes, class reorganisation is desirable, or even inevitable, and it proves impossible or difficult to combine two different modes of teaching in the same class. For various reasons, pupils may have to leave the experimental class. During a period of two years, a considerable number of pupils will actually leave the experimental schools.

One of the main reasons for depletion of the number of experimental children is that teachers often find it convenient to divide the experimental classes into groups—so that they can teach parts of the class by the experimental method and other parts by conventional methods. This loss, in itself, is, from the point of view of the experiment, very unfortunate, but it brings with it a change in the sample which can be even more so, for the teachers will usually divide the class up in a highly selective way. Let us consider this difficulty more closely. Suppose that an appreciable number of teachers decided that only the brighter children should be taught by the Dienes method. This would mean that the group of children taught by this method would, by virtue rather of their talent, than of the method by which they had been taught, stand a good chance of obtaining high marks when, eventually, they were tested. Their performance could mistakenly be interpreted as reflecting the merit of the Dienes method rather than the selection of a specialised sample of children.

To some extent we can cope with the situation in which a teacher uses general ability as a criterion for selecting pupils for tuition by the experimental method, for we have an assessment of this characteristic of each pupil, and, consequently, can take it into account in making our comparison. Unfortunately, we are not so well placed with regard to some other criteria which the teacher may employ in his selection. Since some teachers find that the experimental methods pose problems of discipline, there is a tendency, in certain cases, to choose well-behaved children for tuition by the experimental method. It is not unreasonable to suppose that, during the experimental period, well-behaved children will tend to progress further than others from the level that they had reached at the beginning of the experiment. Since we have no measure of 'goodness of behaviour', we shall not know, in such cases, whether progress is due to the experimental method or to selection of this kind.

Again, one of the experimental methods—the Dienes—involves a great deal of reading, for pupils are required to carry out tasks in accordance with instructions which they must read from work-cards. As a consequence, in selecting pupils for tuition by this method, teachers tend to favour those of high reading-ability. Now, in the case of some of our tests, we have been unable to dispense with written instructions and questions (insight tests are particularly difficult to construct without the inclusion of a considerable amount of written explanation).* Therefore, it is possible that children who are selected for tuition by the Dienes method will do well on certain of our tests—not because the Dienes method is effective, but because these pupils have been selected for their ability to read. Since we have no assessment of the pupils' reading-ability, we have no way of assessing the degree to which, in such cases, this factor has contributed to performance.

Another change of experimental conditions—one that had not been anticipated, and which would have been difficult to control—was due to the employment, in the experimental classes, of special teaching forces. In some cases, the headmaster would take over the experimental class, and give it the benefit of his possibly superior teaching. Just as frequently, he would abstract from the class a special experimental-method group, which he would coach. Sometimes he would join the class-teacher in 'team-teaching' the experimental class. Many reasons can be given for the adoption of this kind of policy—the individual work required by the experimental method needs closer supervision, the assistant teacher is insufficiently acquainted with the new method, the headmaster has a special enthusiasm for it—and so on. Thus, the class would be likely to enjoy teaching of a higher quality, or a heavier teaching force, than would be likely to be accorded to the control classes.

---

*Although these tests are read out to the examinees, the good readers will still be at an advantage, for they will be able to study the questions at their leisure.

How can we cope with the difficulties mentioned in this section? Luckily, an appreciable part of our sample has been unaffected by these disturbances to the design of the experiment, and it should therefore be possible both to make a fair comparison of the teaching methods (although over a considerably smaller sample of pupils than it had originally been hoped) and to make some kind of assessment of the effect of the disturbances.

The problems of assessment

One of the problems involved in the comparison of two teaching methods is that of constructing testing instruments that will favour neither, and yet register the special merits of both. This problem seems to have defeated many of those who have tried to assess structural methods, for, in many cases, conventional arithmetic tests have been used in the assessment, and such tests—insensitive as they are, to the kind of outcome at which structural methods aim—could not have been *expected* to reveal the particular merits of these methods. One could, of course, err in the opposite direction, by constructing tests that would be sensitive *only* to the special merits of structural methods. In this case, one could ensure that the structural methods would shine.

The basic problem, in the construction of a means of comparing structural with conventional methods, is that the teaching objectives of these two kinds of method are not entirely coincident. Our hope for a solution to this problem is justified by the fact that they are *partially* so. For example, there would be little point in constructing a test which would ascertain only the pupil's ability to compute in number systems other than the decimal, for such a test would be bound to favour the Dienes method—just as would be a test of the early mastery of fractions be bound to favour the Cuisenaire method. To compare the Dienes or the Cuisenaire method with conventional methods in such respects as *these*, can be considered to be unfair—for most conventional methods do not aim to develop at an early age the ability to work in many number systems, or to cope with fractions. However, it is possible to construct a fair test, for two reasons: (*a*) the objectives of different methods *do* clearly overlap; (*b*) the agreement of teachers over certain teaching-objectives is so clear in some areas that one is entitled to say that a method that does not achieve these objectives (even where it has not aimed for them) would be deemed inadequate.

An example of a kind of objective on which the proponents of even widely different methods would agree, is that teaching methods should prepare the pupil for the kind of learning that he will be expected to engage in later on in his school work. In this case, one could also say that a method, even if it were not geared to this objective, *ought* to be, and, whether or not it *tried* to prepare the pupil for subsequent learning, if it failed in such preparation, it could be said to have failed as a teaching method. Unfortunately, within the time-span of the present phase of this investigation, we shall not be able to compare methods in terms of this kind of outcome.

We have tried to construct instruments of appraisal which will do justice both to the common aims of the methods under comparison, and to the generally-accepted standards with which early primary teaching is expected to comply. We have constructed tests to measure various aspects of the performance of conventional computing techniques, tests of the ability to apply techniques in the solution of conventional kinds of arithmetical problem, and 'insight' tests. An attitudes questionnaire has also been constructed. It might be pointed out that to some extent, in the case of the problem-arithmetic test, and to a great extent in the case of the insight test, the pupil is expected to cope with situations for which he is unlikely to have been specifically-prepared. In these cases, the particular syllabus-coverage of the methods being compared is likely to have only a remote effect upon performance. Moreover, in the case of the attitudes questionnaire, the effect of syllabus-coverage is likely to be minimal. Although, for the most part, our instruments of appraisal are meant to be insensitive to such differences in performance as would be the direct and obvious result of differences in the syllabuses and aims of the methods to be compared, some provision has been made for the assessment of certain of the less obvious of the consequences of these differences. It would be pointless to compare two methods in terms of aspects of performance for which the syllabus of only one of the methods provided, for the results of such a comparison would be a foregone conclusion. We know already, for example, that Cuisenaire children would have some attainment at computation with fractions, whereas children who had not been taught about fractions would have next to none. We know, likewise, that Dienes children would to some extent be able to compute in number-systems other than the decimal, whereas children whose experience had been confined to the decimal system would seldom be able to venture beyond it. However, there are areas of performance in which differential success, although not a *direct* reflection of differences in the syllabuses and teaching aims of the methods, certainly could be expected to be influenced by them; in these areas, a comparison of methods could reveal differences which would be informative. For example, it would be interesting to know whether acquaintance with a variety of number-systems improves the pupil's ability to deal with denominate number. We have therefore included, in our tests and questionnaire, items which would detect such hypothesised effects of syllabus and aims.

The problems of measuring understanding

For several reasons, it would seem that tests of mathematical understanding would constitute appropriate criteria for the assessment of our methods of teaching arithmetic. In the first place, few teachers would deny that they aim to foster in the pupil an understanding of what he learns—perhaps quite rightly, it is these days taboo *not* to aim to teach for understanding. Secondly, we cannot say that the understanding of a particular notion is merely a reflection of the syllabus covered by a particular teaching-method, for, as it is normally used, the word 'understanding' implies *more* than the ability to reproduce a skill that has been acquired; therefore, we should be gaining some significant information about a method if (perhaps as an *indirect* consequence of its syllabus-coverage) we discovered that it conduced to 'understanding'. Thirdly, since, when we talk of 'understanding', we usually mean, among other things, a grasp of *general* principles, and since general principles are likely to underly a great variety of specific techniques, it is not, perhaps, surprising, that we can find common criteria of understanding which will be relevant to teaching methods with greatly differing teaching-policies and syllabus-coverage.

Although it would seem to be very desirable to compare methods in terms of the understanding resulting from their use, there attach to the measurement of this aspect of attainment many problems which are difficult to surmount. It is relatively easy to define and construct tests of the ability to compute, for this is little more than the ability to perform certain specifiable kinds of operation—an ability which we can test by requiring the performance of a representative sample of such operations. However, as we at present use the term, 'understanding' is more difficult to define, and it is more difficult to know when we are managing to test it, for we have not decided upon the precise behavioural implications of imputing it to a performance, or upon the kinds of achievement that we can take as its token.

Probably, the definition and testing of understanding will follow the course that was taken by those of intelligence: tests *seeming* to divine what *seems* to be what we mean by understanding will be constructed, and those of such tests which eventually prove useful will form the criteria for a behavioural definition of this concept. However, at present there are no such commonly-accepted tests of understanding, so we have had to construct our own.

How can we say what these tests measure? To a certain extent, we can expect this to be self-evident, for the test items will certainly be novel to the pupil, certainly involve the appreciation of mathematical principles (two conditions that can be expected to require the exercise of understanding) and, since they do not rely heavily upon computation, are unlikely to be contaminated by this last aspect of arithmetical performance. As a means of ensuring that these items measured what they appeared to measure, we questioned a sample of pupils after they had completed the test in draft form, in order to ascertain whether the items had been correctly performed by virtue of an understanding or incorrectly performed by virtue of a lack of it. This is in many ways an unsatisfactory basis upon which to lay a claim for the validity of a test, but, at this point in the development of tests of understanding, little more could be done.

Experimental stimulation

Wherever a new teaching-method is to be compared with an established one, there is a possibility that any difference in the results obtained from the two methods is due to the *novelty* of the new method, rather than to qualities that would remain with the method after it, too, had become established. The possibility that this factor has been operating should qualify our acceptance of many of the results of educational research, and the possibility that it will have operated in the case of this particular experiment is a warning that certain of our results must be interpreted with care.

The factors contributing to experimental stimulation are many, and, in some cases, subtle: First, novelty arouses curiosity, and may well be responsible for an appreciable increase in attentiveness to arithmetic on the part of both pupil and teacher. Second, the new circumstances created by a change in method permit the pupil to escape from bad habits: distractions can be *learnt*—very often it is easier to concentrate in new circumstances than in circumstances in which there are well-worn paths along which attention can wander from the object of study; again, failure-roles tend to become attached to particular kinds of achievement, and, since arithmetic is particularly conducive to the development of such roles, it is a subject in which the opportunity to escape from them is likely to prove to be telling. Third, in many schools, pupils have developed a resistance to activities that they are able to see as continuous with traditional school-work. Such activities come to be categorised as 'work' (which is to be avoided), as opposed to 'play' (which is to be embraced). Partly because the new methods involve physical manipulation, they tend to be regarded as play-activities for some time after they have been introduced into a school. However, there are indications that they will eventually become assimilated to the 'work' category, and join the victims of the resistances that operate upon the activities in this category. Fourth, various factors conspire to give the pupil the impression that the work he is doing with a new method is important. Parents become curious, and tend not only to discuss it with their children after school, but to use it as a point of contact with the teachers: sometimes meetings are held for the express purpose of explaining the method to parents. Visitors tend to come to schools using new methods, and, if only out of duty, glow with approval at the pioneering work that is being carried on. Special arrangements are made for group-work, and, as mentioned in an earlier section, the headmaster keeps in close contact with the class using the new method. Fifth, it seems that, in some cases, the new methods constitute an important 'gimmick' with which pupils can oppose and impress adults. In a rapidly-changing culture, children learn at a very early age that their loyalties are divided between identification with adults and identification with peers. Just as they find identity by comparing themselves with their parents, so do they, in some ways, by emphasising the differences. Very frequently the pupil taught by the Dienes method will have parents to whom the workings of number-systems other than the decimal are a complete mystery. In such cases the pupil may develop for the method the kind of enthusiasm that he reserves for activities that are beyond the range of his parents' appreciation (say pop-music) and may at the same time be proud to have been able to reverse the parent/child relationship of expert/learner.

Hitherto we have been considering factors which will affect, primarily, the pupils —but certain aspects of the *teacher's* behaviour are *also* likely to be affected by experimental stimulation. Even where teachers have begun their career with an understanding of the mathematical principles underlying arithmetic, such understandings are likely, after several years of teaching, to have become overlaid with habitual and mechanised ways of conceiving the subject. However, when confronted with a new method, teachers will need to reconceive their arithmetic, and their techniques of teaching will have to be radically overhauled. While this is happening, there will be a tendency to think much more actively about the subject, and to refer much more to basic principles. The teacher's insight will be awakened. Such an awakening is likely to infect the pupil, who is also likely to benefit from the consequences which it will probably have in the enlivening of teaching-technique. Unfortunately, there is no guarantee that the new method will not eventually have had its day, and that the teacher's insight will not go back to sleep. Again, teachers, as well as pupils, tend to get the impression that now arithmetic has become important; they tend also to see themselves as pioneers, and become more strongly involved in their teaching than they normally would be. A consequence of such involvement is that they want the new method to succeed—and this, too, adds to the effort they put into their teaching. Finally, where the teacher has just adopted a new method which has clearly-formulated objectives (as do those methods in question) he is likely to have a much more precise idea of what he is aiming at, and may well teach in a much more streamlined manner, than he will later, when these objectives will have become modified in the direction of his own preferences. Because of this, and because there is probably much less systematic similarity among the objectives of the various control methods, it is probable that teaching with the experimental methods will produce certain characteristic patterns of results which are not likely to remain characteristic, and perhaps will not even remain.

What can be done about experimental stimulation? It has been suggested that one could 'stimulate' the control group, so as to rule out this difference between the two groups. To a certain extent this is possible, and, to the extent that they have been performing yearly tests, we can assume that the pupils in our control group have been stimulated. But we cannot say that the control pupils have been stimulated in the same ways and to the same extent as have the experimental pupils. Under what circumstances *could* we say this? It is difficult to conceive.

A further suggestion, which really amounts to much the same as the foregoing, is that the control group should be taught under 'placebo' conditions—that is, under conditions which replicate all aspects of those obtaining in the experimental group (including novelty) excepting for that aspect which is under investigation. In the present case this would not have been possible, for we could not have found a conventional teaching method which, under any existing teaching circumstances, could have been guaranteed to provide the stimulating impact that the experimental methods are suspected to produce.

We have used three devices for either controlling or measuring experimental stimulation:

First, some of our comparisons will be *between* new methods—the Stern, Cuisenaire and Dienes methods will be compared with one another, as well as being compared with conventional methods. In these comparisons we can more safely assume that the methods being compared will, in relation to experimental stimulation, be on an equal footing. However, even here, there is no excuse for assuming that experimental stimulation will have *no* effect on the results. Perhaps, in the case of some methods, the initial impact is more transient and thus less representative of their lasting effects, than in that of others.

Second, we are assessing the new methods after they have been in use for one year, and again after they have been in use for two years. In this way we hope to gain some indication of the kinds of performance that are being affected by experimental stimulation.

Third, we are comparing experimental with control pupils not merely in terms of overall arithmetical performance, but also in terms of *patterns* of performance over various kinds of task. While most aspects of arithmetic-learning are likely to be affected by experimental stimulation, there may be many that are affected to the same degree. By means of the device mentioned in our previous point, we shall be able to establish which these are, and shall therefore be in a position to assume that certain differences in *pattern* of performance reflect inter-method differences that are not due to experimental stimulation.

THE TEST PERFORMANCE OF 'COLOURED' CHILDREN

by Oscar Ferron

Faculty of Education, Njala University College, Sierra Leone

Editorial Note:

The paper which follows represents a major contribution to our thinking in this country about immigrant children and the reader's attention is earnestly directed to Dr. Ferron's conclusions, especially pp. 55 and 56.

Dr. Ferron is deeply interested in the problem of assimilating immigrant groups and would be happy to be in touch with anyone in this country interested in the problem. Letters addressed to him c/o the N.F.E.R. will be forwarded.

The evaluation of the various races of mankind, and especially comparisons between races of European and non-European stock, has long been the subject of lively controversy. With the invention of intelligence tests in the early part of the century, there emanated a flood of literature on the subject of racial differences for the next thirty years, so that a review of industrial studies would be far too overwhelming to be of much value in the present context. It is sufficient for our purpose to observe that most of the studies up to the early 1930s in the U.S.A. confirmed the superiority of intelligence test performance of:

(*a*) White children over Negro and Red Indian children, test scores of the latter increasing with the proportion of the admixture of ' white blood '.

(*b*) Children of North European Stock over children of South European Stock.

In connection with (*b*) above, the studies by Professor Terman and others seemed to demonstrate the intellectual inferiority of such groups as the Spanish-Mexican, Portuguese, Italian and Slav, the investigators maintaining that: ' . . . these differences are largely independent of educational status, and are to be attributed to innate racial differences '. It is rather interesting to observe, however, that in such comparisons, the performance of the average Negro child compares favourably with that of immigrant European groups. Thus, for example, in one rather typical study (Hirsch 1926), the following means were obtained:

	Mean I.Q.
British Canadians	93.8
French	85.3
Negroes	84.6
Portuguese	82.7

In fact, in most studies, the performance of Negro children has been found to be superior to those of Spanish-Mexican descent.

North American Indians

As far back as 1926, Helmer devised tests which contained subject matter relating closely to the life of various Indian tribes represented at the Haskell Institute. The tests were of the 'best answer' and other standardized types, but contained references to Indian religious ceremonies, social customs and food. On many of these tests the white children were definitely inferior. In 1941 Dennis experimenting with 152 Hopi children aged 6-10 years, obtained a mean I.Q. of 108.3 on the Draw-a-man test.

On the other hand, several studies conducted with American Indians have indicated that the mean I.Q.s for various tribes fall between 69 and 88. However, Rohrer has called the authenticity of some of these results into question on the grounds of:

(*a*) Differences in social and economic status
(*b*) Differences in cultural environment
(*c*) Language differences
(*d*) Differences in test motivation and psychological rapport, and
(*e*) Inaccuracies in estimating chronological age.

After an extensive review, he found that in each study two or more of the above differences could have been said to have contributed to the results obtained. He therefore decided to experiment with a group of Osage Indians, who were at the time (1941), socially, educationally and economically on a par with the white population in the district, English being the language used by all, except the oldest members of the tribe. He compared 125 Osage children in grades I-III with a comparable group of 125 white children, using the Goodenough Draw-a-Man test, and obtained mean I.Q.s of 103.8 and 102.9 respectively, with standard deviations of 21.2 and 19.9. Using the Otis scale with two other groups of 110 Osage children and 110 white children in grades IV-VIII, he obtained mean I.Q.s of 100.1 (S.D. = 18) and 98.1 (S.D. = 17.7) respectively. Although the above figures suggest a slight superiority in favour of the Osage children, neither of the differences was in fact statistically significant. Furthermore, contrary to the results of previous investigations, he found no correlation between the degree of Osage blood and test intelligence.

American Orientals

One of the first investigations in this field was that by Pyle in the early part of the present century. Pyle found that Chinese children attending local public schools were superior to American children in rote learning, but inferior in tests involving logical relationships, and in speed of learning generally. Combining the results from his battery of tests, he found that the efficiency of the Chinese boys was only 84% that of American boys, and Chinese girls 77% that of American girls. The main criticism levelled against this study was the use of unstandardized tests. Later, Yeung, using standardized tests, tested 105 Chinese children aged 9-11 years selected at random from the oriental schools at San Francisco. He obtained a median I.Q. of 97, which compared favourably with that of white children living in the same city, if we make some allowances for the fact that Chinese children must have suffered from a cultural and linguistic handicap.

In the same decade (1926) Darsie found that the median I.Q. of 10-15 year old Japanese children living in large cities in California was 99 when tested on the Binet scale.

In the ' Army Beta ' which is a non-verbal test, the Japanese children were actually found to be superior to the white children with whom they were compared. In another investigation conducted by Sandiford and Kerr of the University of Toronto, 500 Chinese and Japanese pupils attending the public elementary schools of Vancouver city were tested individually by means of the Pinter-Paterson Performance Scale, which obviates the use of language. It was found that the median I.Q. of the Japanese boys was 115.4, of the Japanese girls 112.8, of boys and girls together 114.2. The median I.Q. of the Chinese boys was found to be 107.8, of the Chinese girls 107, and of both together 107.4. The authors concluded that the superiority was undoubtedly due to selective emigration, those ' . . . Japanese and Chinese possessing the qualities of cleverness, resourcefulness and courage . . . ' having emigrated, while the dullards were left at home. This is a moot point, for it is also possible to argue that those who are comfortably established in their native land do not normally have sufficient incentive to emigrate. On the other hand, while the above figures do suggest an element of selection, selection might also have been a factor in the inferior test performance of certain European and coloured immigrant groups investigated. The fact of the matter seems to be that intelligence test scores being a function of both hereditary and environmental factors, and in the absence of evidence relating to innate racial inferiority, one should expect, as Professor Porteus pointed out way back in 1924 that: ' . . . Japanese, Chinese and American children educated together in American schools reach approximately the same scores in intelligence tests . . . ' In this connection research with immigrant children attending primary and secondary schools in Britain might prove to be rather revealing.

Hawaiian groups

The fact that white and coloured children in the U.S.A. live in the same area and attend similar schools, does not necessarily guarantee that their environmental circumstances are similar. The existence of colour prejudice and other social pressures may also affect intelligence test scores, as a result of differences in emotional and motivational tendencies. In Hawaii the environmental factor can be somewhat more effectively controlled. Chinese, Japanese, Portuguese, Puerto Ricans and Filipinos, all originally came to the islands as plantation labourers. The whites and the indigenous people have probably exerted their influence in the past, but they are now less important numerically than the other groups. The Chinese, Japanese and Portuguese have, to a much larger extent than the other races, broken away from the plantations, and now form the bulk of the present Hawaiian population. It has been suggested that as all groups initially had the same social and industrial status and none enjoyed special social and educational privileges, their present socio-economic level is largely the result of their own efforts and inborn capacities. But this is a mere suggestion and a somewhat dubious one at that, for it is impossible to unravel the matrix of perceptual and motivational tendencies that are based in culture and tradition. However, as a result of some seven years' research, in 1930 the University of Hawaii, Honolulu, published the results of testing nearly 4,500 children in Hawaii and on the mainland of America. The tests used were the Stanford-Binet, Porteuz Mazes, and form and assembling tests. Four groups of children were selected as follows:

I. One group from the mainland was made up of 1,590 children who were either native born Americans, or those of Jewish or Italian ancestry. All the children were in attendance at elementary or high schools in Vineland, New Jersey. This group will be referred to as the ' Vineland Whites '.

II. The second group consisted of 363 white children attending Punahou Academy in Honolulu (' Punahou Whites ').

III. The third group was made up of 556 white children belonging to the ' superior middle class ' at Lincoln School, whose socio-economic status as a group was somewhat lower than the Punahou group.

IV. The fourth group of 1,961 children consisted of Japanese, Chinese, Koreans, Portuguese, Hawaiians, and others of mixed ancestry, in attendance at various schools all over Hawaii.

The following table shows the combined results of the tests expressed as test quotients:

Gp. No.	*Ethnic Origin*	*N*	*Mean I.Q.*
I	Vineland Whites	1,590	96.7
II	Punahou Whites	363	111.2
III	Lincoln Whites	556	114.3
IV	*Mixed Group*		
	(*a*) Japanese	419	104.7
	(*b*) Chinese	284	101.5
	(*c*) Koreans	67	103.3
	(*d*) Portuguese	288	93.0
	(*e*) Hawaiians	171	93.4
	(*f*) White Hawaiians	466	102.7
	(*g*) Chinese Hawaiians	266	98.5

Unfortunately, we have no means of ascertaining whether or not these differences are statistically significant. However, as different aspects of intelligence were assessed by means of different tests, and the results were combined to give a single index, they do suggest that when the differences in environmental opportunities as between children in advanced civilizations, and those in under-developed areas respectively, are not too great, the differences in intelligence test performance can be expected to be correspondingly smaller. Moreover, Porteus assures us that the difference of 11.7 points between the mean scores of the Japanese and Portuguese respectively, must surely be significant, and he points out that: '. . . the theory that white blood inevitably confers superiority, at least in the capacities measured by the tests, receives a decided set-back as a result of this investigation '.

Other studies in Hawaii like those of Louttit in 1931 confirmed the main findings of Porteus' investigation described above, but Leiter's study deserved special mention, because he endeavoured to construct a culture-free test, using common and familiar materials. His ' International Performance Scale ', which consisted of matching devices, picture completion, colour forms and the like, was standardized on a group of 1,430 Chinese and Japanese children in Hawaii, between the ages of 6 and 17 years. Its reliability coefficient was found to be .91, and correlations with the Stanford-Binet and Porteus Maze tests worked out at .79 and .71 respectively. The mean I.Q. for the Japanese was found to be 101.3 with a standard deviation of 15, for the Chinese 96.5 with a standard deviation of 14.6. One hundred and forty six Chinese-Hawaiian children obtained a mean I.Q. of 98, while 98 Caucasian children at the 6, 9 and 12-year old age levels scored means of 114, 115 and 98 respectively. The mean for the Caucasian group as a whole works out at 109.

Fijian groups

The colony of Fiji consists of 250 islands, 80 of which are inhabited. According to the 1936 census, the population was made up as follows:

Full-blooded Fijians	97,651
East Indians	85,002
Europeans (i.e. British, Australian and New Zealanders)	4,028
Mixed European and Fijian	4,574

plus a small number of other Pacific islanders and Asiatics. The indigenous population is made up of Polynesians, a tall copper coloured race with oriental features, and the Melanesians, a shorter, much darker race with typical negroid features. The Indians were brought over in 1897 to supply labour for the sugar-cane plantations, and more than 70% of the Indian population in 1936 were Fijian born. The ' Europeans ' tended to be a highly selected group engaged in big business, administration and professional work, and consequently enjoyed the highest socio-economic status of all groups. In 1935 Mann administered his ' Fiji Test of General Ability ' to more than 4,000 school children in Fiji. This test is a picture performance scale based on the local environment, and thus obviates the use of language. A Reliability Coefficient of .85 was obtained, and Validity Coefficients using:

(*a*) the estimates of 4 reliable teachers
(*b*) the Goodenough Draw-a-Man test (750 Fijians)
(*c*) the Otis Self Administering Test (Intermediate form—250 Europeans)

worked out at .77, .78 and .74 respectively. An analysis of the test scores yielded the following main results:

(*i*) The mean scores were found to be in the following descending order: Europeans, Half-castes, Indians and Fijians. All differences were statistically significant.

(*ii*) The differences in the mean scores for the four age groups 10:6, 11:6, 12:6 and 13:6 were statistically significant for the Fijians and Indians, but not for the Europeans.

(iii) There were no significant differences between the mean scores of the older European and Half-caste groups, although when the groups were compared age for age, the differences were statistically significant in favour of the Europeans.

Malayans and Ceylonese

Lewis' investigation in 1949 was a comparative study of the intelligence and educability of Malays and Chinese in Malaya. Using a group test based on geometrical designs devised by himself, as well as Raven's Matrix test, he finally concluded: '. . . in general, Chinese pupils, age for age, have shown themselves to be superior to Malays in test intelligence. Further examination, however, reveals that while this superiority undoubtedly exists, especially in the lower age groups, it is decreasing as the pupils grow older'. The following factors were tentatively suggested as having some bearing on this observed differential:

(*a*) The greater competitiveness of the Chinese, who are considered 'foreigners' in Malaya.

(*b*) The fact that many Malay children transfer from the vernacular schools to English Schools. This is not so in the case of the Chinese who either continue in their vernacular schools or begin their education in English schools.

(*c*) The rural character of the Malays as contrasted with the urban character of the Chinese.

(*d*) A much higher mean socio-economic level among the Chinese (excluding Malay royalty and nobility).

(*e*) Even a differential incidence of malnutrition and hookworm amongst Malay and Chinese respectively was put forward as a factor which might have resulted in differences in mental ability and in motivation.

In 1954, Strauss administered the California Test of Mental Maturity to 16 sub-cultural groups of students in the University of Ceylon. He found that the performance of the Ceylonese students was superior to that of American university entrants on the verbal tests, but inferior on the non-verbal tests, although for 'Total Mental Factors', there was little difference. He attributed these results to the following factors in Ceylonese society:

(*a*) The traditional attitudes towards manual work and manipulative activities.

(*b*) The emphasis on memory work and rote learning.

(*c*) The higher prestige accorded to verbal scholarship.

(*d*) '. . . the reinforcement of certain of these Eastern culture traits as a result of the long period of British rule . . .'

Only 2 of the 16 sub-cultural groups showed a tendency to deviate from the above pattern, but this tendency was not significant even at the .05 level. Another surprising result was that the Burghers, a group of mixed European and Ceylonese descent, had an average performance that was 'much higher' than the average performance of the group as a whole, as well as the various sub-groups.

Strauss also conducted a pilot study in the low country village of Pelpola, with a group of 26 boys and 20 girls, to whom he administered the primary series of the California Test of Mental Maturity. He found that these village children did not compare favourably with the American standardization population. However, of all the test sections, the second highest score was obtained for vocabulary, while the lowest score was for logical reasoning. While the disparity as between verbal and non-verbal test performance for this group was not as pronounced as for the university students, it did exist. This is rather significant, because the trend is in the opposite direction to what one would expect on the basis of rural-urban differences in Europe and the U.S.A. Strauss concluded that: '... There are grounds for believing that these low scores are a direct reflection on the mechanical and doctrinaire type of education to which these children are subjected, and which is itself a reflection of certain basic thought-ways of this culture ...'.

West Indians

In a recent investigation (1965) reported by Vernon, a random sample of 100 English boys aged about 11 years was compared with a similar group of 50 West Indian boys drawn in groups of 10 from 5 contrasted primary schools in the West Indies. The following tests were used: Vernon's Graded Arithmetic-Mathematics Test (first 40 items with a 20-minute time limit), Reid's (1961) English Attainment Test specially constructed for use in the West Indies, the Mill Hill Junior Form A Vocabulary Test, Terman Merril Form LM Vocabulary, a word memory, an information memory and a concept formation test, together with a battery of tests consisting of 31 items representing a wide range of developmental concepts described in Piaget's writings. The Piaget battery consisted of: (*a*) Time Concepts; (*b*) Left and Right; (*c*) Equidistant Counters; (*d*) Logical Inclusion; (*e*) Visualizing Water in a Tilted Bottle (*f*) Conservation of Volume of a Liquid; (*g*) Conservation of Amount of Plasticine; (*h*) Visualizing Insect; (*i*) Number Concepts; (*j*) Conservation of Length of Rods; (*k*) Transferring Dot to Reversed Sheet of Paper; (*l*) Position and Size of Shadow; (*m*) Conservation of Area.

West Indian percentile scores were expressed as quotients on English norms with a standard deviation of 15. The median quotient for Arithmetic was found to be 84, which according to Vernon represents a retardation of 1½ years, which is not too serious when one bears in mind that the median length of schooling of West Indian boys is at least 1½ years shorter than that of the top class in an English primary school. On the English attainment test, the median quotient was found to be 82, but it is rather significant that on the spelling sub-test the quotient rises to 94. Once again in the Word Memory test which measures an ability which can be drilled, it is as high as 91, but drops to 72 for 'Information Memory', which involves meaningful material. In the Concept Formation tests a median quotient of 90 indicates that the gap between British and West Indian performance is not too serious. In this connection Vernon remarks: 'Previous workers, either using concrete objects (Price Williams, 1962) or Kohs Blocks (Jahoda, 1956) have likewise found that retardation of intellectual development in certain pre-literate cultures cannot be identified with Goldstein's inability to abstract '. When the scores in the various sub-tests comprising the Piaget battery were combined, the West Indian quotient worked out at 86, which indicates only a moderate degree of retardation. In comparison with the English scores, the largest differences were found in:

(*i*) Number Concepts
(*f*) Conservation of Volume of Liquid
(*j*) Conservation of Length of Rods
(*m*) Conservation of Area
(*h*) Visualizing Insect

while they are negligible in:

(*g*) Conservation of Amount of Plasticine
(*d*) Logical Inclusion

and quite small in:

(*a*) Time concepts
(*c*) Equidistant Counters
(*e*) Visualizing Water in a Tilted Bottle.

Factor analysis revealed three distinct components which Vernon has designated:

F I .. Numerical orientational
F II .. .. Conservation and
F III .. .. .. Visualization.

Factor I appears to comprise abilities measured by the numerical items, orientation of handedness and time. Factor II loads all the conservation items including some time concepts and ' logical inclusion ', while Factor III involves at the visualization items. From the scores obtained, it would appear that West Indian boys perform relatively well in the orientational aspects of Factor I although they are retarded in number comprehension. They seem to be weak in all the visualizing items in Factor III, except in (*e*) Visualizing Water in a Tilted Bottle which gave the smallest difference in comparison with English scores. They tend to be most backward in Factor II, ' (*g*) Conservation of the Amount of Plasticine ' being an exception. It is suggested that the ability measured by this item may be well developed because of opportunities to play with earth and clay. The publication of further results is now awaited.

African groups

In 1929 Fick compared the performance of Zulu children of 10-14 years of age with that of white South African children using the American Army Beta tests. He found that an average of only 1.2% of the Zulus reached or exceeded the medians of the whites at similar ages. Later in 1939, in his book, *The Educability of the South African Native*, he wrote ' . . . around the ages of 13 and 14 native children are from 4 to 5 years inferior to European children in educability as gauged by the results of intelligence tests . . . . The inferiority occurring in certain tests in which learning or environmental conditions are equalized for the native and European groups does not appear to be of a temporary nature '.

Needless to say, Fick has been severely criticized in this contention by several subsequent investigators. His work was obviously biased, and controls were grossly inadequate.

In 1932, Oliver began developing his ' General Intelligence Test for Africans ', a picture group test which was standardized on several hundred East African children. Unfortunately, however, the work of establishing norms was never completed. With the standardization sample, the test produced a reliability coefficient of .8 and a validity of .6. In one of his studies, he tested 93 African schoolboys of average age 19:6, and 124 European schoolboys of average age 15:0. ' Only 14% of the Africans reached or exceeded the European mean and the African mean was 85% of the European '. Experience with African school children as well as objective research evidence suggests that very often age operates as a disadvantage rather than as an advantage, and it is quite common to find that the youngest children in the class are the most intelligent. However, due to the extreme difficulty of controlling the various culture and environmental variables involved, unlike Fick, Oliver refrained from making racial comparisons.

In 1934, Porteus applied Leiter's International Performance Scale to 197 natives all over 15 years of age, and belonging to different South African tribes. The mean mental age for the whole group was found to be 10 years, the difference between ' mission ' and ' raw ' natives being a little more than one year of mental age. This study has been criticised by Mann on the grounds that numbers were small, and the groups could not be regarded as being representative. Moreover, he doubted the international value of the Leiter scale, which was standardized on Chinese and Japanese children in Hawaii. He felt that he could not accept the contention that the African level of performance was equivalent to that of 10-year-old oriental children in Hawaii.

In another investigation, using his ' Mazes ' test on groups of South African and Australian natives, Porteus thought he had some evidence for the superiority of the Australian ' Blackfellow '. However, the 1936 Australian census showed 53,698 full-blooded aborigines and 23,461 half-casts, and Porteus admitted that only 200 natives could be tested. These were mostly in the mission stations, so that it is possible that the Australian group was not as representative as the South African group. Mann has produced further evidence to suggest that the observed difference is due more to factors arising from contact with and assimilation of white culture than a genuine racial one. He found that by rearrangement of Porteus' data, the mean test age for 128 Australians was 10.89, for the 207 South African natives 10.27, which gave a critical ratio of only 1.8. On the other hand, for 108 mission schooled natives, both Australian and South African, the mean test age was found to be 11.26, while for 127 unschooled natives of both races, it was 9.27 which gave a critical ratio of 6.

Biesheuvel who is considered one of the greatest authorities in this field, did a good deal of experimentation in the forties and fifties. In one such experiment undertaken in Johannesburg he compared a group of 125 native school chidren with a similar group of 125 white South Africans. However, although the two groups were matched for age and sex, they were grossly dissimilar in educational status. The difference between the means was more than one school ' standard ', but Biesheuvel maintained that as the standard of instruction in the native schools was so poor, the effective difference was probably about three years of schooling. From the point of view of socio-economic status, the African sample was, if anything, superior to the mean for the native urban population. The European sample, on the other hand, although superior to the African group, was well below the average for the white Johannesburg Community as a whole, their mean on the South African Group Intelligence Test being 97. However, quite apart from the observed differences in the material environment which favoured the whites, there were also more subtle differences from the point of view of enrivonmental stimulation in the home and neighbourhood background, and the inter-personal relationships between parents and children, which could have affected the African test scores adversely. The tests used were Koh's Blocks, Cube Construction, Alexander's Passalong test, Porteus' Mazes, and a Match test which was specially constructed, being a performance version of tests like number series and figure classification. The language media used were both English and Afrikaans, the children being sufficiently conversant with both these languages to understand the instructions given. The following were the main findings:

(*a*) There were no significant differences between the means on the Passalong, Maze and Match tests, the critical ratios being .4, ·3 and .6 respectively. In all three tests however, the scatter of African scores was much greater than that of European scores.

(*b*) In the Koh's Blocks and Cube Construction tests, the African mean was significantly lower than the European mean, the critical ratios being 2.97 and 3.8 respectively.

Biesheuvel therefore concluded that '. . . Africans are inferior in the ability to manipulate spatial relations, particularly in three dimensions . . . this inferiority has a genetic basis. Yet the possibility cannot be excluded that the difference would vanish if African children had more opportunity in their homes to get acquainted with the conventional spatial arrangements which are an essential part of our daily activities'. Of course Vernon's finding (see p. 48) partly confirms this view, and experience with African students and testees would add further weight to it, but the present author maintains that this deficiency is probably no more genetic than the European's disability in balancing feats *vis-à-vis* the African, which is such a common feature of daily life in Africa (see also McFie).

During the last decade, there have been a few investigations in this field in various parts of Africa, notably those of King, Dunstan, Dawson and Ferron in Sierra Leone, Schwarz in N. Nigeria, Nwana in the Western and Anowi in the Eastern Regions of Nigeria respectively. The investigations by Brimble and Irvine in Rhodesia, Lule in East Africa, Scott in the Sudan, McFie in Uganda and Levine in Nigeria, are also worthy of mention. Apart from the fact that some of these investigations have not yet been completed and reports on others have not yet become available in Africa, it would be impossible to summarize the results of every investigation in this field. We are mainly concerned with the most significant ones and those that have some relevance for educational work in Africa. Accordingly we shall confine ourselves for the present to the investigations by McFie (1954), King (1957) and Dunstan (1961), and report certain results obtained by the present author (1964).

In 1954, McFie tested some 26 new entrants, aged between 16-19 years, to a technical school. The boys, who were in the main from the Ganda tribe, had completed the full primary course and had had from one to three years in a secondary school. In the technical school they studied such subjects as building, carpentry, motor maintenance, machine fitting, etc. After two years they were re-tested and it was found that the small gains they made in Comprehension, Similarities, Arithmetic and Picture Arrangement were not significant. However, statistically significant differences were obtained in tests relating to:

(*a*) Picture Description and Memory for Design

(*b*) Block Designs.

The author finally concluded that, allowing for individual differences, the ability to perceive visual material as wholes (Gestalt perception) is no more innate than the ability to deal with verbal or arithmetical problems. He went on to say: '. . . As long as constructional toys and other geometrical tasks do not occur in their environment, their abilities in these respects will remain undeveloped and they will be handicapped in their ability to deal with the representational and mechanical features of European culture '.

In 1956, King devised the ' Leone Intelligence Test ', and used it at the Sierra Leone Common Entrance examination that year. The test gave a correlation of over .9 with the Schonell Essential Intelligence Test. With 489 ' Colony ' (now Western Area) and 374 ' Protectorate ' (now the Provinces) children, he obtained mean I.Q.s of 80 + and 70 + respectively. The investigator was strongly of the opinion that this was an indication of a general retardation of the child population, and attributed it mainly to environmental factors, such as illiteracy among parents, chiefly mothers, and extremely low standards among teachers in the primary schools.

Dunstan experimented with over 400 Form I pupils in certain secondary schools in Freetown, and analysed the results of the following tests:

(1) Moray House Test of Intelligence (M.H.T. 44)
(2) ,, ,, ,, ,, English (M.H.E. 20)
(3) ,, ,, ,, ,, Arithmetic (M.H.A. 20)
(4) Thurstone's Primary Mental Abilities Test for ages 7-11 years (P.M.A.)

The following is a summary of the main findings:

(*a*) The low mean score on the spatial relations test (Thurstone) confirmed Biesheuvel's results in regard to Koh's Blocks and the Cube Construction tests (see p. 51). In addition, Dunstan found no correlation with age, and the girls' performance was even worse than the boys'.

(*b*) The P.M.A. Vw and the M.H.T. 44 tests gave the same mean mental age of 10:6, showing that the level of performance of these selected (by the Common Entrance examination) Freetown secondary school children is about the same as that of children 3 years younger in the standardization sample.

(*c*) The P.M.A. 'N' test showed a retardation less than for the verbal tests.

(*d*) The most disturbing of the low scores were those for two reasoning tests (P.M.A.) Rw and Rf, in which a mean mental age of 10:0 was obtained. Rw uses words, Rf diagrams. It was found that the mean Rf score was significantly better than the mean Rw score, the excess being equivalent to about 6 months of mental age. This would suggest a language handicap.

(*e*) While ' lack of specific information required for the task ' was found to be an important source of error, an even more important source was ' inadequate formulation of the task '. There was a tendency to deal with a related (and to the child, presumably easier) problem than with the one set.

(*f*) ' . . . there is no evidence of systematic variation of ability in test performance with occupational status of the parents, of the children tested '. According to Dunstan, not only do the children of all occupational classes go to the same primary schools, but also the psychological circumstances of the home environment do not appear to be significantly different.

(*g*) The author suggested that an important contributory factor in limiting the performance of the children was probably the quality of the primary education to which they had been subjected.

Previous investigators in Africa have criticized the use of western type intelligence tests with African children owing to the language handicap. The research evidence does suggest that the gap between white and non-white children can be reduced by the use of non-language tests, but there is also evidence to suggest that where cultural ways of life and patterns of child rearing are significantly different, or because of cultural mores and traditional attitudes, African children may be even more seriously handicapped in picture and performance tests. Theoretically it should be possible to devise an intelligence test that is suitable for any particular culture, but such a test would be useless for purposes of inter-cultural comparison. Moreover, if this were done in certain West African cultures, whose chief status symbols are now essentially western, then the school-going sections of those communities would be at as great a disadvantage with a test based on the native culture as they are with tests standardized in Britain and America.

Provided no invidious comparisons are to be made between cultures, it would seem to be far more constructive to continue with the use of western type tests, since African aspirations are in the direction of a western technological civilization. Comparison between British and African test performance can also be justified on the grounds that such comparisons not only help us to see where the emphasis ought to be placed in the matter of environmental deficiencies, but also they should lead to the improvement of our tests to make them more suitable for Africans. Thus, in an investigation which was started in Freetown in 1960 and completed in Zaria in 1964, the present writer used the ' Leone ' test devised by an African for African children. Analysis of the test results revealed how unsuitable even such a test can be for African children. On the other hand, the material that has resulted has thrown some light on cultural procedures, classroom orientation and test construction, and given a certain amount of direction for future action. One part of the investigation was concerned with comparing several groups of school-going children in Sierra Leone and Nigeria with similar groups in English schools, to obtain the following test quotients:

(1)	English Grammar School (boys)	117	(N= 50)
(2)	Secondary Modern 'A' Stream (girls)	103	(N= 34)
(3)	Secondary Modern 'B' Stream (boys)	96	(N= 36)
(4)	Secondary Modern 'C' Stream (boys)	82	(N= 18)
(5)	Freetown Grammar Schools (boys)	91	(N= 50)
(6)	Freetown Grammar Schools (girls)	95	(N= 50)
(7)	Mean Sierra Leone 'Colony' I.Q. (see p. 52)	80 +	(N=489)
(8)	Mean Sierra Leone 'Protectorate' I.Q.	70 +	(N=374)
(9)	Zaria Grammar Schools (boys)	81	(N=190)
(10)	Mission Primary Schools, Zaria	74	(N=380)
(11)	Native Administration (Muslim) Primary Schools, Zaria	66	(N=100)

In addition to those that are obviously so, the following differences were found to be statistically significant: (3) and (5), (4) and (5), (5) and (6) and (10) and (11). There are in fact good grounds for believing that the test quotient is to a large extent dependent upon the degree of assimilation of western culture.

Tribal comparisons were also made in Zaria in respect of four main groups: (*a*) Yoruba, (*b*) Hausa/Fulani (*c*) Ibo and (*d*) Miscellaneous tribes. The only statistically significant differences that were found were those between the Yorubas and the other groups. The reason for this was probably because the Yorubas had the highest percentage of highly selected secondary school pupils and no native administration pupils in the group considered. On comparing the performance of 32 Christian Yorubas with 26 Christian Hausa/Fulani in those Christian schools where, obviously they would have been exposed to the same kind of environmental stimulation, it was found that the Hausa/Fulani were superior, the difference in mental age of three months being significant at the P =.2 level only.

Conclusion

An examination of the evidence reviewed above would suggest that, except perhaps in the case of selective mating and selective migration, there is no concentration of what Hebb refers to as ' Intelligence A ' in any particular geographical region of the world. That differences in ' Intelligence B ' and ' Intelligence C ' exist between various racial groups there can be no doubt, but these differences are by no means confined to Europeans on the one hand and non-Europeans on the other. Such differences are also found to exist within coloured and white groups respectively, depending upon the nature of the tests used. The evidence also indicates that the test performance of coloured children can not only equal, but also surpass that of white children. On the other hand, it can also be disconcertingly low when compared with European standards. But of one thing we are now reasonably certain and that is that test performance is not a function of the degree of 'white' ancestry. One is inclined to feel that it is much more closely related to:

(*a*) the degree of assimilation of 'white' culture, and
(*b*) the emotional and motivational tendencies that are based on culture, tradition and survival.

As far as one can make out, poor test performance of coloured children is fundamentally the outcome of the emotional and motivational problems involved in the assimilation of western culture. The difficulty of passing from one world to another as the child grows older, of living in two worlds simultaneously, and of acquiring the newly imposed and rather unrealistic symbols of success and status, represents a struggle for survival with which the biological struggle of the survival of the fittest can bear no comparison, for even those who survive the former appear to be emotionally, psychologically and intellectually maimed for life.

Where circumstances are such as to ensure that white and coloured groups share a common way of life and have equal educational opportunities, differences are small or non-existent. Examples of this are found in the case of the American Negroes as compared with certain immigrant European groups, the Osage Indians and Orientals in the U.S.A. and Hawaii. The test performance of half-caste groups the world over also tends to be on a par with that of Europeans of the same socio-economic level, provided the emotional problems of the former are not too intense. In the main the test performance of half-caste children tends to be superior to that of the indigenous natives, a finding that is confirmed when one compares the performance of mission and 'raw' natives in the Porteus (1934) investigation, the Freetown Creoles with tribal Sierra Leoneans, and the children in Christian schools with those in Muslim schools in Nigeria. By and large the differences within the coloured group can be explained on the basis of both cultural assimilation and the motivational tendencies that are culturally based. Thus in most British and American investigations Jewish children have returned a mean I.Q. higher than British and American children of non-Jewish origin, the Japanese appear to be superior in test performance to the Chinese, the Chinese to the Malays. In the case of groups like the West Indians or the Freetown Creoles certain attitudes towards authoritarianism and rote learning have become crystallized as a result of the speed with which these people were required to assimilate western culture, and the dearth of properly trained teachers that was inevitable.

As far as the work of education in Africa is concerned, the tendency among Europeans to condemn Africans as being inherently inferior in intellectual potentialities has no scientific justification whatever. Expatriates working in Africa must face up to the fact that if African test performance is low when compared with European standards, steps must be taken to make good certain environmental deficiencies. Of course, this constitutes a tremendous undertaking; but unless and until a simultaneous attack is made on many fronts, progress will continue to be slow. It is the present author's contention that progress can be speeded up considerably if, instead of concentrating the bulk of our energies and available finance at the higher levels of the educational pyramid, we distributed our attention more evenly over the whole field of education from birth onwards.

After a mere six-year period in West Africa, the writer can make no claim to an exhaustive study, but he is inclined to the belief that the following are the main contributory factors in the poor test performance of West African children, which judicious planning and realistic administration can quite conceivably control:

A. *The School.* Assuming the school to be the starting point and the teacher the chief agent of change, we might mention the following:
  - (*i*) A mechanical approach to teaching in the early years particularly.
  - (*ii*) The teachers' relative lack of facility in the use of the English language so that communication between teacher and child is not as effective as it might be.
  - (*iii*) discouraging the child from asking questions and exploring his environment.
  - (iv) Severe and capricious corporal punishment particularly in the early years.
  - (*v*) Mental rigidity and an attitude of dogmatic authoritarianism.

B. *The Home.* If we have some success with the items outlined above, then the pupils of today, who will become the parents and teachers of tomorrow, will be in a position to make a contribution towards bringing the following under reasonable control:
  - (*i*) Birth injuries and the incidence of subsequent malnutrition and disease.
  - (*ii*) Certain early child rearing practices that are in the nature of traumatic experiences.
  - (*iii*) The nature of the emotional relationships between adults and children, and the relative absence of an atmosphere of approval and encouragement.
  - (*iv* Illiteracy, mainly of the mother, and a lack of appropriate stimulation in the home and neighbourhood environment.
  - (*v*) Lack of definite teaching on the part of older children and adults.

Governmental action. In addition to their commitments in (A) above, might it not also be reasonably expected that the powers that be will take tangible steps to control the following?
  - (*i*) Lack of proper physical surroundings in home and school.
  - (*ii*) Relative lack of stimulation from mass media of instruction.

A final note. If most of the above environmental variables can be brought under control, as in the case of immigrants in Britain for instance, where colour prejudice and discrimination among children are not as crippling as in the U.S.A., it seems to us that a very useful link and exchange of ideas might be established between research workers in Britain and those in other parts of the world. For our part, we would welcome any confirmation or refutation of research findings which would give us further guidance in helping to improve the test performance of coloured children.

Bibliography

Anastasi, A. (1958). *Differential Psychology.* New York: Macmillan.

Biesheuvel, S. (1943). *South African Intelligence.* South African Institute of Race Relations.

Biesheuvel, S. (1949). 'Psychological Tests and their Application to Non-European Peoples', in: *Year Book of Education.* London: Evans.

Biesheuvel, S. (1952a). 'Personnel selection tests for Africans', *South African Journal of Science*, vol. II, nos. 2 and 3.

Biesheuvel, S. (1952b). 'A study of African ability', *African Studies*, vol. II, nos. 2 and 3.

Biesheuvel, S. (1959). *Race, Culture and Personality.* The Hoernle Memorial Lecture.

Dawson, J. L. (1963). 'Psychological Effects of Social Change in a West African Community'. Unpublished Ph.D. thesis, University of Edinburgh.

Dunstan, G. (1961). 'An Investigation of Problems Relating to Selection for Secondary Education in Sierra Leone'. Unpublished Ph.D. thesis, University of Edinburgh.

Ferron, O. (1964). 'A Study of Certain Factors Relating to the Tested Intelligence of Groups of Some West African Children'. Unpublished Ph.D. thesis, University of London.

Havighurst, R. J., Gunter, M. K., and Pratt, I. E. (1941). 'Environment and the Draw-a-Man Test: the performance of Indian children', *J. Ab. Soc. Psych.*, 41, 50-63.

Hebb, D. O. (1949). *The Organisation of Behaviour.* New York: Wiley.

Hirsch, N. D. M. (1926). 'A study of natio-racial mental differences', *Gen. Psych. Monographs*, 231-406.

Irvine, S. H. (1963). 'Some Practical and Theoretical Problems of General Ability Testing at the African Std.VI Level in S. Rhodesia', *East African Institute of Social Research*, Conference, Part B.

Irvine, S. H. (1964). 'Test Abilities and Attainments in Africa'. Paper read at a meeting of the British Psychological Society, Education Section.

Jahoda, G. (1956). 'Assessment of abstract behaviour in a non-western culture', *J. Ab. Soc. Psych.*, 53, 237-243.

King, V. E. (1957). 'A Study of Individual Differences among Primary School Children in Sierra Leone with special Reference to the Selection of those Children for Secondary Education'. Unpublished Ph.D. thesis, University of London.

Klineberg, O. (1924). 'An experimental study of speed and other factors in racial differences', *Arch. of Psych*, no. 8, 57-74.

Lewis, G. E. D. (1949). 'A comparative study of the intelligence and educability of Malays and Chinese in Malaya and its significance in relation to educational policy, with special reference to the political future of the Malays'. Unpublished Ph.D. thesis, University of London.

MacArthur, R. H., Irvine, S. H., and Brimble, A. R. (1964). *The Northern Rhodesian Mental Ability Survey*, 1963. Lusaka: Rhodes-Livingstone Institute.

McFie, J. (1961). 'The effect of education on African performance on a group of intellectual tests', *Brit. J. Educ Psych.*, 31, 11.

Mann, C. W. (1939). 'A test of general ability in Fiji', *Journal of Gen. Psych.*, 54, 435-454.

Mann, C. W. (1940). 'Mental measurement in primitive communities', *Psych. Bulletin* 37, 366-391.

Mumford, W. B., and Smith, C. E. (1938). 'Racial comparisons and intelligence testing', *Journal of the Royal African Society*, vol. XXXVII, no. CXLVI, Jan. 1958, 46-57.

Notcutt, B. (1950). 'The measurement of Zulu intelligence', *J. Soc. Res.*, vol. 1, no. 2.

Porteus, S. (1924). 'Temperament and mentality in maturity, sex and race', *J. Appl. Psych.*, 8, 57-74.

Porteus, S. (1930). 'Race and social differences in performance tests', *Gen. Psych. Monog.* 8, no. 2.

Porteus, S. (1937). *Primitive Intelligence and Environment.* New York: Macmillan.

Price-Williams, D. R. (1961). 'Analysis of an intelligence test used in the rural areas of Central Nigeria', *Overseas Education*, 33, no. 3, 124-133.

Price-Williams, D. R. (1962). 'Abstract and concrete modes of classification in a primitive society', *Brit. J. Educ. Psych.*, 32, 50-61.

Reid, L. H. E. (1961). 'The Construction and Standardization of Attainment Tests for Secondary School Selection in Jamaica'. Unpublished M.A. Thesis, University of London.

Sandiford, P., and Kerr, R. (1926). 'The Intelligence of Chinese and Japanese children', *J. Educ. Psych.*, 17.

Schwarz, P. A. (1961). *Aptitude Tests for Use in the Developing Nations.* Pittsburg: American Institute for Research.

Silvey, J. (1963). 'Testing Ability Tests'. *East African Institute of Social Research.* Conference, Part B.

Strauss, M. A. (1954). 'Sub-cultural variation in Ceylonese mental ability: A study in national character', *J. Soc. Psych.*, 39, 129-140.

Vernon, P. E. (1960). *Intelligence and Attainment Tests.* London: University of London Press.

Vernon, P. E. (1961a). *The Structure of Human Abilities.* London: Methuen.

Vernon, P. E. (1961b). *Selection for Secondary Education in Jamaica.* Kingston: Govt. Printer.

Vernon, P. E. (1962). 'Intellectual development in non-technological societies', *Proc. Intern. Cong. Appl. Psych.*, 14th, vol. II, 94-105. Copenhagen: Munksgaard.

Vernon, P. E. (1963). *Personality Assessment—A Critical Survey.* London: Methuen.

Vernon, P. E. (1965). 'Environmental handicaps and intellectual development', *Brit. J. Educ., Psych.*, XXXV. Pt. 1.

Research Note:

FURTHER EVIDENCE ON THE EFFECT OF DATE OF BIRTH ON SUBSEQUENT SCHOOL PERFORMANCE

by RICHARD FREYMAN,
Educational Psychologist, Middlesbrough Education Committee

Introduction

THE influence of children's date of birth on their educational progress has recently been studied more closely (Jackson, 1964; Jinks, 1964; Johns, 1962; Williams, 1964). It was found that, generally speaking, a larger proportion of 'B' and 'C' stream than 'A' stream children tended to be born during the period May to August, i.e. towards the end of the scholastic year. This fact lends support to those who doubt the educational and even the moral validity of streaming children according to their ability.

The writer became interested in the subject when studying the date-of-birth distribution both amongst children taught by peripatetic teachers in remedial reading groups and others referred to him in his capacity as educational psychologist, because of reading retardation. Almost half of these children were born during May-August.

Jinks' study in *Educational Research* (June, 1964) came to hand whilst the investigation was under way and provided an interesting means of comparison.

The sample

The children (boys and girls) involved were (*a*) 3,419 children born during 1952-1955 from eight streamed junior schools in different parts of the town; (*b*) 4,120 children born during 1953-1955 from seven streamed secondary modern schools, four grammar schools and two other selective secondary schools; (*c*) 364 junior school children attending remedial reading groups; (*d*) 312 children referred to the child guidance service both for educational and clinical reasons; (*e*) 156 children attending the day Special E.S.N. School. First year children in secondary schools were not included as it was thought that age allowances built into the 11+ selection procedures might distort the results. Streaming in the selective secondary schools was not taken into account, because the methods of class organization used in these schools tended to be different from the usual pattern of streaming children according to ability.

A. Results

In both the total junior school and total secondary school samples slightly more children were born during the May-August period than during either the September-December or January-April periods (Table 1).

Table 1—Dates of Birth (Total Sample)

	Sept.-Dec.		*Jan.-Apr.*		*May-Aug.*		*Total*	
	No.	%	*No.*	%	*No.*	%	*No.*	%
Junior Schools	1,037	30.4	1,178	34.4	1,204	35.2	3,419	100
Secondary Schools	1,309	31.8	1,393	33.8	1,418	34.4	4,120	100
Total	2,346	31.3	2,571	34.1	2,622	34.6	7,539	100

Table 2—Dates of Birth (Months) of Streamed Junior School Children

	Sept.	*Oct.*	*Nov.*	*Dec.*	*Jan.*	*Feb.*	*Mar.*	*Apr.*	*May*	*June*	*Jul.*	*Aug.*	*Totals*
'A'	150	127	133	117	103	115	115	93	103	90	95	88	1,329
'B'	72	63	77	102	113	115	111	117	108	115	108	115	1,216
'C'	45	48	55	48	77	52	88	79	93	85	117	87	874
Totals	267	238	265	267	293	282	314	289	304	290	320	290	3,419

Table 2a—Dates of Birth (Periods) of Streamed Junior School Children

	Sept.-Dec.		*Jan.-Apr.*		*May-Aug.*		*Totals*	
	No.	%	*No.*	%	*No.*	%	*No.*	%
'A' Classes	527	39.7	426	32.1	376	28.2	1,329	100
'B' Classes	314	25.8	456	37.6	446	36.6	1,216	100
'C' Classes (+one D+one E class)	196	22.4	296	33.9	382	43.7	874	100
Totals	1,037	30.3	1,178	34.5	1,204	35.2	3,419	100

Table 2 shows the number of 'A', 'B' and 'C' children born in each month and Table 2a the number and percentages of these children in each of the three periods of the year in the streamed junior schools.

As can be seen, the date-of-birth distribution in each stream varies widely from that of the total sample. The percentage, e.g. of 'B' and 'C' children born during May-August was well above that of 'A' children born during the same period. On the other hand, the percentage of 'A' children born during September-December was well above that of 'C' children. It was found (*a*) that these differences were statistically highly significantly χ^2 (2df)$=56.62$ $P<0.01$ and (*b*) that statistically a highly significant linear relationship existed for all three streams between length of schooling, depending on month of birth, and the percentage of all members in the appropriate stream born in that month. For the 'A' stream the percentage increased with length of schooling, whereas for both 'B' and 'C' streams the trend was reversed. In other words, the nearer a child was born towards the end of the academic year, the smaller appeared his chances of becoming a member of an ‘ A ’ class.

Test of significance of linear relationship

‘ A ’ stream: $t= 6.76$ (10 d.f.) <0.01
‘ B ’ stream: $t= 3.9$ (10 d.f.) <0.01
‘ C ’ stream: $t=-5.32$ (10 d.f.) <0.01

B. Secondary schools

In Middlesbrough, all 11 + children have in the past sat a one-day examination for the purpose of secondary school selection, consisting of a verbal reasoning, an English and an Arithmetic test.* The test scores were combined with scaled teachers' estimates. Borderline candidates were interviewed by a panel. Children not selected for Grammar Schools but of high average and above average performance in the examination were offered places in either a technical or another type of selective secondary school.

For the purpose of this survey, the figures from the four grammar schools and the two other selective schools involved (one for boys and one for girls) were added (Table 3).

Table 3—Dates of Birth in all Secondary Schools

	Sept.-Dec.		*Jan.-Apr.*		*May-Aug.*		*Totals*	
	No.	%	*No.*	%	*No.*	%	*No.*	%
Secondary Modern Schools	651	28.4	814	35.5	827	36.1	2,292	100
Selective Secondary Schools	658	36	579	31.7	591	32.3	1,828	100
Totals	1,309	31.8	1,393	33.8	1,418	34.4	4,120	100

The most interesting fact was the difference between the number of children born during September-December in the secondary modern and selective secondary schools. This difference was statistically highly significant χ^2 (1 d.f.) $=27.06$ $P<0.01$.

Considering the secondary modern sample separately, it was found that again a statistically significant (at the 0.05 level) linear relationship existed for the ‘ B ’ and ‘ C ’ streams between length of schooling, depending on month of birth, and the percentage of all members in the appropriate streams born in that month.

---

*It has been proposed by the Education Committee that from 1966 onwards junior children will proceed to Secondary Schools without selection and be able at 13 + to transfer to High Schools at the wish of their parents and after consultation with their headteachers.

Table 4—Dates of Birth in Streamed Secondary Modern Schools

	Sept.-Dec.		*Jan.-Apr.*		*May-Aug.*		*Totals*	
	No.	%	*No.*	%	*No.*	%	*No.*	%
' A ' classes	197	31.8	229	36.9	194	31.3	620	100
' B ' classes	157	27	213	36.7	211	36.3	581	100
' C ' classes + (D, E, F)	297	27.2	372	34.1	422	38.7	1,091	100
Totals	651	28.4	814	35.5	827	36.1	2,292	100

Test of significance of linear relationship

' B ' stream: $t = 2.35$ (10 d.f.) $<0.05 > 0.01$
' C ' stream: $t = -2.99$ (10 d.f.) $<0.05 > 0.01$

The fact that the ' A ' classes date-of-birth distribution did not obey the ' rule ', i.e. did not display a linear relationship as well, can probably be explained by the mentioned two-tiers selection procedure existing in this authority. This means that with some exceptions the ability of the ' A ' stream children in the secondary modern schools was generally around the average mark, i.e. much nearer the ability of the ' B ' and ' C ' children than could be expected. In spite of this, however, the percentage of ' A ' children born in the September-December period was still higher than that of the ' B ' and ' C ' children.

Although no comparative figures for streaming in the selective secondary schools were available because of varying methods of streaming, e.g. according to subject rather than ability, it was possible to study the results of the yearly examinations in the 2nd, 3rd and 4th form of one of these schools. The interval between the highest and lowest mark (out of hundred) obtained by the children in each form was divided by three, and the number in each sub-group considered to represent ' A ', ' B ' and ' C ' stream children. 38.5 per cent of the high-mark and 39.6 per cent of the average-mark children were born during September-December, whilst only 21.8 per cent of the low-mark children were born during the same period. In reverse, 47.3 per cent of the low-mark children were born during May-August in comparison with 30.1 per cent of the high mark and 37.2 of the average-mark children. Thus a larger proportion of younger than older children obtained low and average marks.

C. Remedial reading groups and Child Guidance Service

Out of 364 junior children taught by peripatetic remedial teachers in small groups in their schools because of reading retardation, 170 (46.7 per cent) were found to be born during May-August, 93 (25.5 per cent) during September-December and 101 (27.8 per cent) in the January-April periods.

Out of 105 children (aged 5-13) referred by schools to the writer because of reading retardation during three years, 55 (52.4 per cent) were born during May-August, i.e. more than were born in the other two periods combined (21.9 per cent for September-December, and 25.7 per cent for January-April).

The picture changed when the number of these children was compared with 207 children (aged 5-16) referred to him mainly for psychological problems. 84 (40.6 per cent) of these were born during May-August, 71 (34.3 per cent) during September-December and 52 (25.1 per cent) during January-April. Whilst still the largest percentage of these children was born during May-August, the difference between this and that of children born during September-December (6.3 per cent) was not as striking as the differences observed between May-August and September-December born children in the remedial group (21.1 per cent) and the group of children referred to the writer for educational retardation (30.5 per cent).

D. Special E.S.N. School

Out of 156 children attending the local E.S.N. school (aged 8-16), 57 children (36.5 per cent) were born May-August, 38 (24.4 per cent) during September-December and 61 (39.1 per cent) during January-April. Whilst the percentage of the September-December born children was distinctly below that of the total sample of children in the ordinary schools, the percentage of the May-August born children was not very much above that of the total May-August percentage (see Table 1).

Discussion

The results of this study are in line with the findings of previous similar investigations (Jackson, 1964; Jinks, 1964; Johns, 1962). When compared with the normally expected monthly distribution of births, relatively more ' C ' and to some extent ' B ' than ' A ' children tended to be born during May-August. Conversely more ' A ' than either ' B ' or ' C ' children tended to be born in the September-December quarter. Statistically significant relationships existed between length of schooling and membership of ' A ', ' B ' and ' C ' classes in the junior school sample.

The significant (0.05 level) linear relationships also existing between length of schooling in the secondary modern schools and the statistically significant difference between number of children born during September—December in selective and other secondary schools confirmed Jinks' view that the ' disadvantage of being born late in the scholastic year is not eliminated by the end of the primary school period and that the streamed schools tend to perpetuate the differences in performances due to date of birth '. Of special interest in this connection were also the examination results obtained from one of the selective secondary schools. Here a larger proportion of high-mark and average-mark children were found to be born during September-December than low-mark children. This might be compared with Jackson's report (1964) of the findings of an inquiry by the Durham Education Authority concerning the relationship between dates of birth and G.C.E. 'O' level results. ' When 218 G.C.E. candidates from seven secondary modern schools were analysed, 101 had birthdays between September 1st and December 31st, and only 48 between May 1st and August 31st '.

The ' disadvantage of being born late in the scholastic year ' was seen most strikingly in the high percentage of May-August born children in the remedial reading group (46.7 per cent) and amongst those referred to the writer because of reading difficulties (52.4 per cent). Babyishness and perceptual immaturity, necessitating a return to the starting poing of learning basic skills, were characteristic features of these children, almost all of whom were members of 'B', 'C' and 'D' classes. Secondary psychological symptoms such as lack of confidence and emotional maladjustment were often displayed as well.

Amongst causes for these difficulties the one mentioned most frequently has been the lack of sufficient time, i.e. three years, spent by May-August born children in the infant school. September-December born children are almost eight years when transferring to their junior schools after having spent three years in the infant schools. May-August born children, however, are only between seven and seven years four months at the time of transfer, having spent only two years or a little more in their infant schools. At the developmental stage of children aged between seven and eight years, a difference of nine to twelve months still represents an appreciable disparity between learning capacities of the younger and the older children within the age group.

It is, of course, true that 28.2 per cent 'A' class junior school and 31.3 per cent 'A' class secondary school children were born during May-August. Similarly 37.4 per cent of the 'high-mark' children in the secondary school examination sample were May-August born. We must obviously look for other factors besides the date-of-entry effect which are responsible for determining which type of May-August child will find itself in a 'C' rather than an 'A' class. One of the most important of these is no doubt the social factor. Its influence has been described fully by both Douglas and Jackson. It could not be dealt with at length in this study.

Of some interest were the figures obtained from the E.S.N. school. The percentage of children (36.5) born during May-August was not very much above that found for the total sample (34.6 per cent) (Table 1). It was also lower than that quoted for E.S.N. children (42 per cent) by Williams. Perhaps Williams' sample contained a number of children of the type that in areas with a well developed remedial reading service might be helped in the ordinary schools, i.e. would not need to attend the E.S.N. school. It must be remembered in this connection that the percentage of May-August children in the remedial groups was 46.7. Maybe the difference between the number of May-August born children in the E.S.N. school on the one hand, and the remedial reading groups on the other hand, indicated that educational sub-normality is a specific condition rather than a delayed mental development, evident in many children born late in the scholastic year.

Conclusion

The evidence accumulated can leave no doubt about the educational, social and psychological disadvantages of a large number of children born during the period May-August. Streaming, as Jackson has shown so admirably, accentuates these disadvantages. More teachers than ever before are becoming aware of this fact and are beginning to think in terms of non-streamed education. A large number of teachers do, of course, still feel that streaming is necessary. The writer hopes that his evidence will contribute to a re-thinking in these teachers, particularly with regard to the fate of the May-August born children. These find themselves often in the lower streams because of an accident of birth and not because of basic lack of ability.

Generally speaking, a new educational deal for these children is required. Many suggestions have been offered. Longer time in the infant school, more ' flexible forms of educational organization ' and ' the extension of services providing guidance for individual children early in their school career ' (Williams) have been mentioned as having an important bearing on this matter.

As far as the ascertainment of E.S.N. children is concerned, great caution, unless other evidence was overwhelming, should be exercised when ascertaining those born during May-August. Retardation in their case might be due to immaturity rather than to a mainly sub-normal mental condition.

Acknowledgement

I wish to express my thanks to Mr. R. Thomas, Senior Lecturer of the Department of Mathematics, Constantine College of Technology, Middlesbrough, for his assistance with the statistical work involved. I wish also to thank the headteachers of the schools involved for their co-operation in supplying their pupils' dates of birth.

References

DOUGLAS, J. W. B. (1964). *The Home and the School.* London: MacGibbon and Kee.
JACKSON, B. (1964). *Streaming: An Education System in Miniature.* London: Routledge and Kegan Paul.
JINKS, P. C. (1964). 'An investigation into the effect of date of birth on subsequent school performance', *Educational Research*, VI, 3, 220-225.
JOHNS, E. M. (1962). 'The age factor in reading retardation', in: *Researches and studies*, University of Leeds Institute of Education, no. 24, 1-7.
WILLIAMS, P. (1964). 'Date of birth, backwardness and educational organisation', *Brit. J. Educ. Psychol.*, 34, 247-255.

Research Note:

MERITOCRATIC AND SOCIAL CLASS SELECTION AT AGE ELEVEN

by D. F. Swift,

Department of Extra-Mural Studies, University of Liverpool

Introduction

THE present debate regarding the organization of education in society would be clearer for greater agreement in the terms used and a better understanding of what the education system is doing for and to society. The following two perspectives upon the human allocation activities of the education system are offered as a preliminary step in research. Adequate information on each of these activities would give us a much sounder factual background to our attempts at reorganization of education. It is not intended that these perspectives should lead to mutually exclusive categories which together cover all that the education system does for society. They are an attempt to arrive at the sort of information which is necessary to an understanding of what education is doing for society at the present time. That is, the questions which current debates are raising have to be answered, to some extent, in these terms.

Analytical perspectives

Meritocratic Selection. The first form of selection will be known as ' meritocratic ' selection. Broadly, this derives from a notion that access to the most important, most prestigeful, and best paid occupations in society should be reserved for the most intelligent. It is generally accepted that, in the developing industrial societies, the educational system is replacing all others as an avenue of social mobility (Floud 1950, Banks 1958, Halsey 1961). Because of this, it is assuming a more and more strategic position for carrying out the function of a sieve of ability, distributing occupational rewards according to its definitions of intellectual merit.

This perspective encourages analysis of the system of education in terms of how it allocates ' ability '. It also helps to focus attention upon the many different possible definitions of ability and their usefulness.

Social Class Selection. Education may also be seen as one of the methods by which the social class system perpetuates itself whilst providing for some social mobility. That is, to the extent that the values and skills of one social class are represented in the normative and behavioural requirements of the school system, children from that social class will be at an advantage. In this way, social mobility is checked.

No judgement need be made at this point about the value for society which a change in social mobility, in either direction, might have. The two extremes are fairly clear. Complete ' openness ' or social fluidity might be disastrous. On the other hand, the predominance of the values and standards of one social class in the educational system might tend to reify these values. This might then work towards reducing the flexibility of the educational system in the face of the changing demands of the economic system. The effect upon the efficiency of the whole society may very well be important.

The aim of this paper is to suggest that two functions of the educational system, meritocratic and social class selection, may be studied with a great deal of benefit to our understanding of the relationships which the educational system bears to the developing society. A simple way in which this may be done can now be described.

Design of the research

The following model was used in a piece of research which the writer conducted in the North of England. The study was primarily concerned with the relationship between parental mobility ideologies and the ability of eleven-year-olds to score on verbal ability and attainment tests. During the course of the research, a group of families (132) was studied in order to compare the attitudinal environment of the home with the ability of the respective children to score on the tests of the 11 + allocation procedure. This group represented a ten per cent sample of all families in two Local Education Authority Divisions and having one child preparing to sit the 11 + examination.

The sample was divided into two further groups according to the social class ranking of the father's occupation. On this basis, 45 families were characterized as ' middle class ' and 87 as ' working class '. (These two groups were basically, non-manual and manual.) The grossest categorization of test performance was a simple split between ' success ' or ' failure ' on the 11 + examination. ' Success ' would mean that the child was to be sent to a grammar school at the expense of the Local Education Authority. ' Failure ' meant that the child would be allocated to a secondary modern school. Each child was represented by his or her actual scores on English, Arithmetic and Verbal Reasoning tests. These formed the testing portion of the allocation procedure recommended by Yates and Pidgeon (1958).

Results

The figures for the entire state-school sampling-frame were as follows:

Table 1

Parental Occupational* Group and Success

(State School Sampling Frame)

	I	*II*	*III*	*IV*	*V*	*VI*	*Total*	*Not Known*
Pass	108	78	62	13	13	4	278	
Fail	98	120	290	44	157	213	922	
	206	198	352	57	170	217	1200	69

These may be collapsed into a table which compares 'social class' performance. After the inclusion of private school pupils the total picture was accurately reproduced in the sample.

Table 2

Social Class and 11 + Success

	Middle class	*Working class*	*Total*
Pass	27 (60%) (73%)	10 (11%) (27%)	37 (100%)
Fail	18 (40%) (19%)	77 (89%) (81%)	95 (100%)
	45 (100%)	87 (100%)	132

Thus we see the middle class child has six chances out of ten of getting to the grammar school under the aegis of the Local Education Authority. On the other hand, the chance of the working class is slightly greater than one in ten. These figures are similar to those found by Floud, Martin and Halsey (1956) in two other areas at the other side of England seven years earlier.

---

*The categories were: I=Professional and managerial; II=Clerical; III=Skilled manual; IV=Foremen; V=Semi-skilled manual; VI=Unskilled manual. A middle class group is formed by amalgamation of Groups I and II.

The next step is to seek the cause of this very powerful imbalance of educational opportunity. This was found in the very different abilities of the social classes in test performance:

Table 3
Sample: Mean Score and Standard Deviations of Test Scores, Grouped by Social Class

	Middle class		*Working class*	
	Mean	S.D.	Mean	S.D.
Verbal Reasoning	117	10.69	100	15.45
English	122	14.21	103	15.63
Arithmetic	117	12.75	101	15.57

When borderlines are drawn up for each of the three tests, the middle class children who come into any one of them are much more likely to have been successful in the overall procedure.* Thus a summary of the ‘instances of class representation’ gives a rather startling picture.

Table 4
Sample: Total Social Class Representation on the Three 11 + Test ‘Borderlines’

	Middle class	*Skilled*	*Unskilled*	*Total*
Pass	30	11	0	41
Fail	9	32	7	48
	39	43	7	89

---

*Each unofficial borderline comprised those children whose test scores fell between the lowest successful score and the highest unsuccessful score. In the case of the verbal reasoning borderline this involved a range of ten points (from 112 to 122).

However, an analysis of the borderline performance of each child represented in Table 4 reveals that there was no evidence of bias on the part of the panel which makes the final decision. The imbalance must be explained by the higher means and narrower dispersions of the distributions of middle class scores.

The conclusion, therefore, is that the disparity between ‘class chances’ is not due to any other factor than differences in ability to score on the tests. Thus, from one point of view, the British ideal of ‘equality of opportunity’ is being met. Access to the first step towards a middle class occupation depends upon the ‘ability’ of the children rather than their social origins. They are making the first step towards the economic elite on ‘meritocratic’ criteria.

Measurement of Meritocratic Selection. However, as most psychologists would argue, the definitions of ‘ability’ themselves tend to have strong social class connotations. We must therefore press the meritocratic notion to a more refined meaning. For Michael Young MERIT=I.Q.+EFFORT (1961). The second term was added because current psychological knowledge recognizes the importance of motivation in test performance. For the purposes of this exposition ‘pure’ merit will be intelligence, and in order to facilitate analysis, the concepts of intelligence, which were suggested by Vernon after Hebb, will be used. Vernon (1957) differentiates conceptually, between intelligence A, which is ‘some quality of the central nervous system ultimately determined by the genes’*; this will be called innate intelligence. Intelligence B represents the exhibition in everyday behaviour of this potential, innate intelligence. Finally, intelligence C may be seen as the samplings of IB which are contained in intelligence tests; this is represented by intelligence test scores.

We are now in a position to consider the results on the intelligence test borderline. They should give us some indication of the meritocratic functions of the selection procedure. Analysis of these scores indicates that there is a bias in the remainder of the selection procedure which supports middle class children who come into the ‘verbal reasoning’ borderline against their ‘intellectual equivalents’ from the working class. Of 14 middle class children in this ‘intelligence’ borderline, only 3 failed; in the working class 10 out of 12 failed.

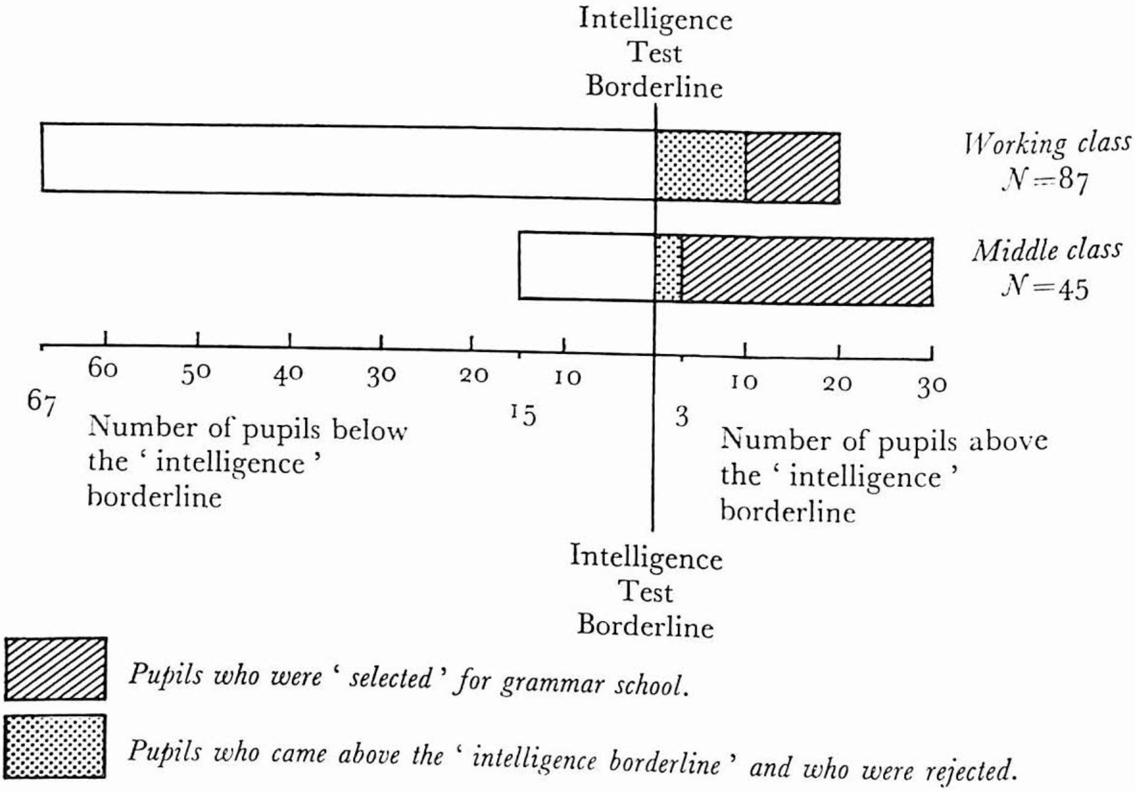
At first sight this may be dismissed as only a superficial distortion to the idea of the school as the means by which ‘bright’ children are given an opportunity to leave the working class. Therefore, to decide on the degree to which the meritocratic function is being performed, these figures must be compared with the numbers of children who actually do get the opportunity. For this purpose, it will be assumed that children in the borderline and above are of ‘similar’ tested intelligence. This has been depicted in figure I.

---

*As such, it is clearly not measurable; however, for the purpose of this paper the use of the concept ‘social class IA mean’ will imply that such a thing does exist.

Figure I

A Comparison between Working Class and Middle Class Success on the 11 + Taking into Consideration Similarity of ' Measured Intelligence '



Thus we have two groups of children—20 working class and 30 middle class—of similar high tested intelligence. Of these 50 children, 13 are to be rejected for grammar school. The result is that one half of the working class children, and one tenth of the middle class children are rejected. Putting it another and more important way, one half of the top twenty per cent of the working class are rejected in comparison with one tenth of the top sixty-six per cent in the middle class.

The purpose of this article is to suggest that there is a way in which the ' meritocratic ' function of selection at any point in time may be estimated or described.

1. Assume that the two social classes have equal innate intelligence means.
2. Assume that the middle class successful proportion indicates the proportion of children who might legitimately (i.e. who are potentially capable) take this preliminary step towards the intellectual elite.*

---

*i.e. ignoring the possibility that the middle class proportion might also be increased.

On this criterion, 27 out of 45 middle class children and 10 out of 87 from the working class are given the chance. Therefore, we can say that the meritocratic function of selection at this point is operating at the

$$\frac{\frac{37}{132}}{\frac{27}{45}} \times 100 = 46.7\% \text{ level.}$$

This means that forty-seven per cent of those children whom we consider to be adequately endowed intellectually are gaining access to to this first step towards elite positions in society.

Until similar data are collected for other years and in other areas, there is almost nothing with which the figure of forty-seven per cent may be compared. We do not know whether it is rising or falling over the years, although we do *hope* that it is rising. However, there is one figure with which it may be compared—100%—and the comparison will be highly relevant to the ‘ scraping the barrel ’ or ‘ more means poorer ’ point of view which some educators take towards opening up British elite education to a few more people. It is relevant, that is, to the tentative plumbing of ‘ the pool of ability ’ which is going on at present.

However, at least one of these assumptions is a highly contentious one. To assume that the social classes have equal innate intellectual means is not realistic in the face of the conventional wisdom of educational psychology. Relaxing this assumption removes all sense from the index as it was calculated.

A further index can be devised which concentrates only upon the children above the ‘ verbal reasoning ’ borderline. The initial assumptions here will be:

1. The children above the borderline are of ‘ similar ’ IC in that they are considered adequate to the demands of the grammar school and hence are potential members of the intellectual elite.
2. Assume that social class is a meaningful criterion for grouping the intelligence scores of children. In this case, the rationale is that because of the similarities between the value-orientations of middle class and school life, the use of middle class children as a base group is justified. Intelligence tests are better tests on middle class children.

Looking only at the borderline children, we can say that the meritocratic function of the 11 + examination is represented by the ratio of the proportion of successful borderline children to the proportion of successful *middle class* borderline children, i.e.—

$$\frac{\frac{37}{50}}{\frac{27}{30}} \times 100 = 66.6\%$$

By this we mean that sixty-seven per cent of all those children who are judged on the criterion of an intelligence test, to be adequate for grammar school education, are admitted to it. This produces a satisfyingly larger percentage, but as with the first meritocratic calculation, attention must be drawn to the assumptions. In this case we would wish to look further into the first assumption of similarity in tested intelligence between the two social class borderline groups. It is reasonable to argue, that in line with all psychological and sociological knowledge, borderline groups which have been chosen on the basis of tested intelligence will produce a working class group of *higher* innate potential intelligence than that of the middle class, whose social ' advantages ' are more likely to allow performance at a level commensurate with their potential. We can assume that the members of the working class group are more likely to be under-performing in relation to their innate potential. Thus, the percentage figure arrived at on these assumptions is overestimating the strength of the meritocratic function in the working class.* On the other hand, the earlier index may provide an understatement of meritocratic selection across both classes because it assumes random distribution of innate intelligence through the social classes. Which figure we accept, or how the percentages might be amalgamated, will depend upon how we can see the assumptions upon which they are based.

The Measurement of Social Class Selection. The function of selection as a stabilizer of social class and an inhibitor of social change may be determined in a similar fashion. This may be shown as the ratio between working class and middle class success, i.e.

$$\frac{11\%}{60\%} \times 100 = 18.3$$

Thus we can say that if 100 represents access to middle class occupations for children of working class parents at the level enjoyed by middle class children, and 0 to be no access whatsoever, then the social mobility function of selection at this point is represented by the index score 18.

Conclusion

Clearly, these percentages will depend entirely on the assumptions underlying the analysis and there are no logical grounds for comparing them with each other. However, it seems possible to compare the indices for the same function at different points in the educational system, and over a period of years, thus producing a ' flow-diagram ' indicating the changing performance of the educational system. It must also be remembered that if a piece of research is to provide worthwhile information upon which discussion about the organization of our talent-producing institutions may be based, it would require much larger numbers in many different areas.

This sort of analysis will depend for its value upon the importance which the educational system has for social mobility, and here a caveat must be added. It is important to note that the perspectives were concerned with the contribution of *education* to allocation in society. If the level of analysis were focussed upon what is happening in the total society on the two criteria of selection, a necessary first step would be to establish how far education was important in the total process of meritocratic or social class selection.

---

*Remembering assumption 1. However, it might also be argued that the above borderline groups cover a sufficiently broad range of intellectual ability for the differential selection to be a function of the shapes of different distributions of the scores.

References

Banks, Olive (1958). 'Social Mobility and the English Educational System,' *International Review of Educatio* vol. IV, no. 2.

Floud, Jean (1950). 'Education and Social Mobility,' *Yearbook of Education*, Ch. V. London: Evans Bros.

Floud, J., Martin, F. and Halsey, A. (1959). *Social Class and Education Opportunity*. London: Heinemann.

Halsey, A. H., Floud, J. and Anderson, C. A. eds. (1961). *Education, Economy and Society*, p. 2. Glencoe, Illinoi The Free Press.

Vernon, P. E. (1957). *Secondary School Selection*. London: Methuen.

Yates, A. and Pidgeon, D. A. (1957). *Admission to Grammar Schools*. London: Newnes, for the N.F.E.R.

Young, Michael (1961). *The Rise of the Meritocracy*. Middlesex: Penguin Books.

Research Note:

DIFFERENCES IN MATHEMATICAL ACHIEVEMENT RELATED TO TYPES OF SECONDARY SCHOOLS AND THEIR GEOGRAPHICAL LOCATIONS

by W. H. King

University of London Institute of Education

Introduction

THE research reported is part of a much larger study* carried out between 1960 and 1963 on comparative standards of achievement between 14-year-old children in South-East England and in the twin cities of St. Paul—Minneapolis, Minnesota. The American children were in the ninth grade, i.e., 14-years old.

The sample

A stratified sampling technique was used to pick out 58 secondary schools so that each school system was represented by type (grammar, modern, technical, comprehensive); by geographical area (London, outside London); by school population; single-sex or co-educational (Table 1).

Table 1

Number of Schools in the Sample: Type, Size and Location

Area	Sex	Type of schools				
		Modern	*Grammar*	*Technical*	*Comp.*	*Total*
London	*All boys*	2	2			4
	All girls	2	2			4
	Co-ed.	4	2			6
	Subtotal	8 (477)	6 (354)			14 (831)
Suburban (large schools)	*All boys*	4	3			7
	All girls	4	3			7
	Co-ed.	4	2			6
	Subtotal	12 (576)	8 (514)			20 (1090)
Suburban (small schools)	*All boys*	2	2			4
	All girls	2	2			4
	Co-ed.	2	2			4
	Subtotal	6 (394)	6 (364)			12 (758)
Total		26 (1447)	20 (1232)	6** (340)	6** (317)	58 (3336)

---

*Final results are not yet available. The main statistical analysis of the results was supervised by Dr. E. Paul Torrance, Director of the Bureau of Educational Research, University of Minnesota, Dr. R. A. Anderson and Professor R. H. Beck of the University of Minnesota, Professor R. T. Johnson of Rutgers State University, Professor P. E. Vernon and the writer, University of London Institute of Education.

**Because of the relatively small number of students involved each of these 6 schools were drawn randomly from the whole area.

Table 2 shows the percentages of students in each category in the sample as well as the proportions of all students in the various categories compared with the total secondary school population in South-East England.

Table 2

Area		Types of Schools			
		Modern	*Grammar*	*Technical*	*Comprehensive*
London	Population	12.64	5·30		
	Sample	14·30	10.58		
Suburban (large Schools)	Population	36.14	18.46		
	Sample	17·32	15·44		
Suburban (small schools)	Population	9·11	2.58		
	Sample	11·75	10.89		
Total	Population	57·89	26.34	6.31	9·40
	Sample	43·37	36.91	10.22	9·50

All figures are Percentages

In spite of a carefully applied random sampling technique, some categories are over and others under-represented. The grammar schools have greater representation since there are 26.34% of children in grammar schools in South-East England, but 36.91% were tested in the sample; consequently 43·37% of the children in the modern schools were included in the sample instead of the 57.89% that might have been expected in the total school population. Again, by using stratified sampling, the unit of testing was a complete form and therefore not all children were in the age-range 14.0 to 14.11 at the time of testing. In fact, 38.0% were under 14 years and 1.6% were over 15 years of age. In addition, there are many obvious pitfalls in classifying schools by a single typed name; the mean scores of some modern schools were higher than those of some grammar and technical schools, although mean scores of all grammar, all technical, all comprehensive schools were probably representative of achievement of these various types of schools.

Class means (using the mean score obtained by a single form as the unit), and the mean scores of all students gave similar results—49.81 and 50.24 for the grammar schools, 44.18 and 44.29 for the modern schools—all means based on scores on Raven's Progressive Matrices.

Tests used

In the complete survey, three standardised tests were used in addition to the teachers' questionnaires and lengthy students' questionnaires. Only the results of the standardised tests are reported now. These tests were:

(*a*) Mathematics Test I by MacFarlane Smith issued by the N.F.E.R. Time: 65 minutes.

(*b*) The Co-operative Sequential Tests of Educational Progress—Mathematics STEP Test. Form 3A—prepared by the Educational Testing Service, Princeton, New Jersey, 1957.

(*c*) Raven's Progressive Matrices, 1956 edition. This is a well-known test of non-verbal reasoning. Its chief weakness in this research was its lack of a high enough ceiling to differentiate those of high ability who scored in the range 50 to 60.

The published age-range of the N.F.E.R. Mathematics Test 1 is given as 12.0 to 13.11, but the range scores obtained by students who were mainly a year older was reasonable. The scores were in the range 5 to 99, with a mean score of 69.99. To some extent this imposed a ceiling effect on the bright children's scores, but not to such a great extent as the Raven's Progressive Matrices. The fact that the contents of this test seemed eminently suitable, and that there was a lack of suitable mathematical tests for adolescents at the time this research was planned, determined to a large extent the choice of this test.

Return of scores

3,134 out of 3,336 students (93.9%) had scores on the three tests. The number of pupils who took individual tests are given in Table 3.

Table 3

Means and Variances for Complete and Incomplete Test Scores

Group	*Raven's P.M.*			*STEP 3A (Maths.)*			*MATHS. 1*		
	N	$\overline{X}$	S^2	*N*	$\overline{X}$	S^2	*N*	$\overline{X}$	S^2
Complete	3134	47.30	52.08	3134	30.69	96.50	3134	69.99	509.48
Incomplete	117	44.13	82.73	132	24.20	89.90	74	53.52	633.29
t or F		4.61**	1.59**		7.44**	1.07		6.19**	1.24*

The t-test was used to test significance of differences of the means and the F-test for differences of variances.
*indicates significance at .01 level.
**indicates significance at .05 level.

The mean score for the 'completes' (i.e. those students with the three test scores) was greater than the mean for the 'incompletes' in all three tests and the differences in the means were highly significant, i.e. could have happened by chance in less than one in one hundred cases. The scores of the 'incompletes' were more variable than the 'completes' for the Raven's Progressive Matrices and for the N.F.E.R. Mathematics 1 Test, but not for the STEP 3A. These results show quite clearly that the omission of these absentees' scores introduced some bias into the analysis of the 'complete' scores.

Comparison of achievement by types of school

Table 4 shows the mean score and variance on each of the three tests for the four types of secondary school. The statistics are calculated on form means and not individual scores, i.e., each school provides a mean from each of the two forms drawn in the sample, so that the mean score on any of the tests taken by the technical schools will be based on twelve means.

Table 4

Test Means and Variances for the Three Tests Classified by Type of School

School	*No.*	*Raven's P.M.* $\bar{X}$	*Raven's P.M.* S^2	*STEP 3A (Maths)* $\bar{X}$	*STEP 3A (Maths)* S^2	*MATHS. 1* $\bar{X}$	*MATHS. 1* S^2
Grammar	20	49.81	0.09	37.13	0.22	83.83	0.53
Technical	6	49.72	0.99	33.44	1.14	80.28	15.83
Modern	26	44.18	0.24	24.22	0.28	55.69	1.25
Comprehensive	6	45.76	1.39	26.62	7.41	57.96	12.42
Total	58	47.18	0.08	30.51	0.68	69.60	0.71
Adjusted means		46.13		28.41		64.83	

The means in the bottom row show the means computed on the basis of the population proportions and not the sample proportions (Table 2), and the fall in the mean scores in all three tests is mainly due to the oversampling of the grammar school population. The same rank order of the four types of schools occurs in the three tests.

The next table shows the significance (on the ' t '-test) of the differences between these mean scores.

Table 5

Tests of Significance for Contrasting Pairs of Schools by Type on each of the Three Tests

Contrasting Pairs of Schools	*Raven's P.M.* Mean	*Raven's P.M.* Mean Difference	*Raven's P.M.* t	*STEP 3A (Maths.)* Mean	*STEP 3A (Maths.)* Mean Difference	*STEP 3A (Maths.)* t	*MATHS. I* Mean	*MATHS. I* Mean Difference	*MATHS. I* t
GRAMMAR with	49.81			37.13			83.83		
Technical	49.72	0.09	0.08	33.44	3.69	3.18*	80.28	3.55	0.87
Modern	44.18	5.63	9.81**	24.22	12.91	18.39**	55.69	28.14	21.11**
Comprehensive	45.76	4.05	3.33	26.62	10.51	3.81	57.96	25.87	7.19**
TECHNICAL with	49.72			33.44			80.28		
Modern	44.18	5.54	4.99**	24.22	9.22	7.75**	55.69	24.59	5.95*
Comprehensive	45.76	3.96	2.57	26.62	6.82	2.34	57.96	22.32	4.32*
MODERN with	44.18			24.22			55.69		
Comprehensive	45.76	1.58	1.24	26.62	2.40	0.87	57.96	2.27	0.61

*indicates significant differences at .05 level.
**indicates significant differences at .01 level.

The results are consistent throughout the three tests. On Raven's Progressive Matrices, grammar schools obtain higher scores than the other three school types, but only the mean difference between the grammar school scores and the modern school scores is highly significant. On the STEP 3A, the grammar schools obtain higher scores than the other three types, significantly different at 0.05 level for the technical schools, and highly significant (.01 level) for the modern schools.

A similar pattern occurs with the mean scores of the N.F.E.R. Mathematics I The grammar schools have the highest mean score on this test and the differences of their means with those of the modern and comprehensive schools are highly significant. The technical schools also obtained significantly higher scores (at 0.05 level) than the modern and comprehensive schools on this test.

In all these test results, the grammar school students are the least variable, i.e. act as a fairly uniform group within a narrower range of ability as measured by these tests; with the other three school types, modern, technical and comprehensive, variability increased in that order.

In summary, the initial differences in ability between school types as measured by the Matrices occurred again in the differences, many significant, on both of the mathematics tests. This must be an artefact of secondary selection. The analysis of between school mean differences revealed consistent superiority of the grammar school in both ability and achievement. Differences in achievement on the two mathematics tests became much more exaggerated when initial effects of differences of ability, measured by the Progressive Matrices, were removed by using an analysis of covariance. These differences related to school types may be due to many interacting causes such as motivation, level of aspiration, teaching ability, etc., in addition to selection. It seems clear from a consideration of the wide range of mean scores of the 58 schools (92.4 to 23.5 on the N.F.E.R. Mathematics Test) and the overlapping of mean scores of different types of schools that a composite mean score from several types of school may have little value and less relevance as a measure of outcomes over a complete school system or of a country's overall ability and makes comparative studies even less worthwhile if a comparison of mean scores is the sole aim.

Comparison of performance of high and low ability groups

The results obtained by students of high and low ability were next considered. High ability students were those who obtained a score of 55 or more, low ability students, those who obtained 40 or less on the Progressive Matrices.

Two-thirds of the high ability group came from the grammar schools, whereas three-quarters of the low ability group are in modern schools. Mean scores and variances were also computed for both ability groups on the three tests.

The pattern of scores for these two groups reproduce the rank order already found between the types of schools for their separate scores and variances for their complete sample populations on the two mathematics tests, but not on the Progressive Matrices. No difference was found between schools, by using analysis of variance, on the Progressive Matrices on either the high or low ability groups, but at both these levels, there was a highly significant difference (.01 level) between types of schools on the two mathematics tests. From the grammar school sample of 20 schools, 253 were rated high ability students and 61 students from 26 modern schools were similarly rated. If it is held that the Progressive Matrices is a reasonable and acceptable measure of mental reasoning ability, then much overlapping must exist in the range of ability found in these different types of school at the age of fourteen years, and no one type of school has a monopoly of students in the high or low ability category. When analysis of co-variance, which allowed for differences in ability as measured by scores on Progressive Matrices was applied to each of the ability groups, it was found that mathematical achievement was, however, still associated with the type of school.

Geographical locations of schools

The 58 schools in the sample were all situated in South-East England, either in or outside London. Any division of this number of schools into London or suburban schools must be made in a somewhat arbitrary fashion. In this research, schools within the area of the London County Council (in 1959, 1960), were classified as London Schools.

Table 1 shows the number and type of schools involved. The mean scores of the schools, based on form means from individual students' scores, were calculated together with the standard deviations. On the three tests, the same rank order is preserved—large suburban schools, small suburban schools and the London schools with the latter consistently showing the least variation on the three tests. Comparisons of mean scores, similar to Table 5 between the three categories of schools are given in Table 6.

Table 6

Tests of Significance for Contrasting Pairs of Schools by Geographic Area on each of the Three Tests

Contrasting Pairs of Schools	*Raven's P.M.*			*STEP 3A (Maths.)*			*MATHS. I*		
	Mean	Mean Difference	t	Mean	Mean Difference	t	Mean	Mean Difference	t
LONDON with	45.82			28.85			66.28		
Suburban large	48.13	2.31	2.92*	32.96	4.11	4.45**	75.39	9.11	5.39**
Suburban small	47.04	1.22	2.79*	30.22	1.37	1.95	67.61	1.33	0.93
SUBURBAN large with	48.13			32.96			75.39		
Suburban small	47.04	1.09	1.34	30.22	2.74	2.97*	67.61	7.78	4.09**

The large and small suburban schools obtain larger mean scores on the Progressive Matrices—both significant differences. The large suburban schools obtain significantly higher mean scores than the London schools on both the STEP 3A and the N.F.E.R. Mathematics Test I; the smaller suburban schools, irrespective also whether they are grammar or technical schools, obtain higher scores than the London schools though the differences in these cases are not significant. There are significant differences between large and small suburban schools on the STEP 3A (.05 level) and the N.F.E.R. Mathematics Test I (.01 level) and although the large schools have higher mean scores on the Progressive Matrices, this difference is not significant.

Conclusions

In spite of the great care with the school sampling there was some discrepancy between the number of students in the obtained and expected distributions. A further effect on the mean scores was due to absent students who had missing scores on one or more of the three tests; their scores tended to show lower ability and attainment on the test scores available. In addition 39.6% of the sample population was outside the age range 14.0 to 14.11 years, although they were all in the same year group in their schools. The third-year students drawn in the sample of schools did not yield as homogeneous a sample of 14-year-old children as had been expected. This, in itself, was not as important in the present report if the sampling technique yielded representative groups of children irrespective of age who were in their third year of secondary school education.

Because of their selected entry and the nature of the subsequent teaching, students from the grammar and technical schools achieved consistently higher mean scores on all three tests. The observed mean differences in achievement on the two mathematics tests were greater than the mean differences in ability, as measured by the Progressive Matrices, between these schools. Analysis of co-variance showed these differences by types of school to be highly significant.

Analysis of the students' scores arranged on their results on the Progressive Matrices into high and low ability groups, showed again the superior achievement of the grammar and technical sample; these differences were even greater when statistical adjustment was made for individual differences in scores on the ability test (Progressive Matrices). In fact, in the major research, of which this report is only a part, the grammar school consistently obtained significantly higher scores (.01 level) than those obtained by the American Public schools. As regards methods of comparison and analysis, the mathematical achievements of students from large suburban schools (over five hundred students) were higher than those in the other two categories.

The results emphasised the overlapping of scores in the different types of schools with a few modern schools obtaining higher mean scores on some of the tests than a few of the grammar and technical schools. The difference between the mean scores of grammar, technical schools and the modern schools was to be expected, but the very wide gap between their achievements and those of the modern schools deserves more urgent consideration since the causes, motivation, level of aspiration, etc., are not solely linked with differences of ability. The major problem to be found in this research is to raise the achievements of the lower ability group in the modern and comprehensive schools without lowering the high standards of achievement reached by students in the grammar and technical schools. The standards of the comprehensive schools (1960/62) as measured by the two Mathematics tests were disappointing; one would have hoped that their mean scores were nearer to those of the grammar schools.

VOLUME VIII, No. 2 February, 1966

EDUCATIONAL RESEARCH

A Review for Teachers and all concerned with Progress in Education

Published for
THE NATIONAL FOUNDATION FOR EDUCATIONAL RESEARCH IN ENGLAND AND WALES
by
NEWNES EDUCATIONAL PUBLISHING CO. LTD.
Tower House, Southampton Street, London, W.C.2

EDUCATIONAL RESEARCH AND THE TEACHER

In a recent editorial* on research communication, we mentioned the work of the Committee to promote 'More Interest in Educational Research' (M.I.E.R.) which had been considering ways in which research might be brought into closer relationship with educational administration, policy-making and classroom practice. A preliminary announcement of a Conference to be organized by the Committee was also made.

The Conference—on Educational Research and the Teacher—was held on 16th October, 1965, at the University of London Institute of Education, and aroused considerable interest. Some 350 people attended, and the Conference was representative of many groups of educationists: administrators; H.M. Inspectorate; teachers in schools, colleges of education and universities; research workers; and editors of educational journals.

The principal speakers were Professor Brian Foss (University of London Institute of Education); Dr. M. Kellmer Pringle, Director of the National Bureau for Co-operation in Child Care; Mr. D. Gratton, Head of Further Education, B.B.C. Television; Mr. P. Taylor, of the Schools Council; and Mr. E. Potter, Headmaster of Forest Hill Comprehensive School. Group discussions were a feature of the Conference and were concerned with the participation of teachers in research; the communication of research findings; the place of research in Colleges of Education, and in courses and in-service training; attitudes to research; and the determination of areas in which research seemed to be needed.

The Conference report, which has been written by Mr. W. van der Eyken, editor of *New Education*, will be published in mid-February. It will be a readable and concise account, presenting the main themes which emerged from the Conference, and a summary of the principal suggestions which were put forward. The report will also include sections listing some of the organizations concerned with research in this country, and some of the more important research projects in progress. An annotated list of periodicals which publish reports on research projects, and a short bibliography of useful introductory books on educational research and research methods, will also be included.

The report—*Educational Research and the Teacher*—is available, price 3s. 0d. including postage, from the Senior Administrator, Publications Section, National Foundation for Educational Research, The Mere, Upton Park, Slough, Bucks.

***Educational Research*, Vol. VII, No. 3, June 1965, p. 164.

SOCIAL CLASS AND ACHIEVEMENT MOTIVATION

by D. F. Swift

Staff Tutor in Social Studies, University of Liverpool

THE basic facts of social class performance in school are so well known as hardly to need repeating. As all teachers know, the children who do the best work, are easiest to control and stimulate, make the best prefects, stay at school longest, take part in extra-curricular activities, finish school with the best qualifications and references and get into the best jobs, tend to come from the middle class. The relationship exists in the Junior School (Douglas, 1964), manifests itself in the 11 + (Floud, 1956; Fraser, 1959; Swift, 1965) and seriously affects secondary education (*Early Leaving*, 1954).

Swift (1965) found that the children of middle class parents had six times as good a chance of selection at 11 + as working class children. In this study it became clear that the lack of balance in these ' class chances ' was mainly due to the fact that the sons of middle class parents were, to that degree, better able to score on intelligence and attainment tests. Many psychologists explain this by suggesting that there is a basic intellectual superiority in the middle class which is maintained by selective mating and a tendency to inherit innate intelligence. This explanation begs as many questions as it answers, but their discussion is irrelevant to the material with which this paper will deal, since all parties would agree that there is room for some environmental distortion of the basic genetic relationships.

A recent study by a psychologist found that all the factors which are usually thought to be bad for the intellectual development of the child (except a working mother) were in fact related to poor performance in the school. Fraser (op. cit.) found that parental education and reading habits, income, occupation and living space all related significantly with I.Q. and school performance. In the motivational sphere, parental attitudes towards the education and future employment of the child, together with the degree of encouragement which they felt they offered the child towards school work, were all significantly related with I.Q. and performance. Finally, abnormality in the family was found to be an important correlate of poor performance.

Of the eleven ' aspects ' of the environment which Fraser considered, the nine mentioned above can be fairly called important elements of the social class situation. Of the remaining two, one (mother at work) was not significant, whilst the second (abnormal background) is to some extent a social class definition.

Since these ' aspects ' are also important elements of the social class system a question immediately arises as to whether they would maintain their significance if the population were separated into groups representing different culture-patterns. Under these circumstances, would the ' aspects ' retain their significance for performance, or does it derive from the fact that they are simply dimensions of a certain social class sub-culture? It is unfortunate that social class itself has not received the scientific ' mapping out ' necessary to its being treated as a variable which can be strictly controlled, and it is possible to suggest that it never will, but this does not make it permissible to ignore it.

The sociological stream of thought which has provided the most valuable theory concerning the relationship between socio-cultural learning and the development of intellectual and motivational skills is the ' clash of values ' approach which began with Allison Davis' (1948) inaugural lecture on social class influences upon learning. (See also: Davis, 1949 and 1960.)

It was Davis' contention that there are barriers to understanding *within* societies which are every bit as serious as those which exist *between* societies; barriers which are all the more serious for the fact that the participating members do not realize their existence. Consequently, Davis enjoined educationists to study the social class system to discover those mental problems, solutions and cultural definitions of what is pleasant and desirable which are emphasized within the different social classes. Similarly, Toby (1957) analysed the problems involved in the maladjustment which occurs when the lower class child brings the attitudes towards education of his culture into the school.

Child-rearing techniques, motivation and school achievement

The Davis critique of objective and subjective intelligence testing in the educational system stimulated research studies which can be grossly separated according to their concern for either child-rearing practices or parental value-systems. These two aspects of the learning process are clearly interdependent and to concentrate in research upon one aspect without the other is to run the risk of discovering irrelevant relationships since important factors in the other perspective are left uncontrolled. This becomes more important the ' nearer ' the research approaches the individual. That is, the more detailed the definition of influential factors the greater the danger that its context is ignored. Child-rearing research has suffered greatly from this weakness and despite valiant efforts to take a synoptic view of the many non-comparable and occasionally contradictory findings (Bronfenbrenner, 1958; Clausen, 1963), headway has been slow. Research has concentrated upon a few of the more obvious mechanisms by which parents attempt to socialize children, and in doing this it has often simplified to a caricature the complex relationship between the environment and the developing personality.

Perhaps the grossest simplification undertaken by these studies is to assume that parental perceptions of what they do to a child are adequate representations of what *is* done to a child by the parent. Human interaction is far more subtle than this. A child is able to ' tune into ' the implicit culture of his family life without too much interference from the manifest attitudes of parents towards what he should do in a particular situation. The folk-lore adage, ' do as I say, not as I do ', comes nearer to an understanding of this complex relationship than many of the researches.

As an example of the contradictory conclusions which these studies have produced, we can compare Drews (1957) with Fraser (op. cit.). Drews discovered the attitudes of mothers of high and low academic achievers of both gifted and average intelligence on scales of permissiveness, protectiveness and domination. Her conclusions were that mothers of high achievers were more authoritarian and restrictive than mothers of low achievers. Parents of high achievers of gifted intelligence also seemed to have more punitive attitudes with respect to child-rearing. In contradiction, Fraser's research implies that the friendliness and spontaneity of a lenient ' democratic ' atmosphere provided the ' best ' (in terms of school motivation and achievement) family environment. Similarly, Elder (1965) investigated the association between perceived parent-adolescent relations and the likelihood of reaching secondary school in the United States, Great Britain, Western Germany, Italy and Mexico. He found a negative association between parental dominance and access to secondary education in all five countries where the structure of educational opportunity made it feasible.

Despite the contradictions in research findings, Bronfenbrenner, by dint of imaginative understanding and a not altogether persuasive hypothesis* perceived a change taking place during the twenty-five year period of research which he reviewed. For Bronfenbrenner, middle class parents in America have changed from a ' restrictive ' to a ' permissive ' approach. Fortunately, he avoids the superficially reasonable conclusion that since there has been no change in middle-class ability to out-perform and out-drive the working class, child-rearing methods are of no importance. As Kohn (1963) points out, the researches summarized by Bronfenbrenner were only as good as the questions they used and hence their implied conceptual models. The changes which took place represent faddish alterations in superficial techniques which are not symptomatic of changes in the basic relationships between parents and children. ' Parents have changed techniques in the service of much the same values ' (Kohn 1963, p. 473).

A further point of view has been brought to bear on the question by Veroff (1965) who suggests a theoretical background to the problem in motivational research of distinguishing between situational and personality factors in any situation. Veroff's major hypothesis is:

> 'that there is a critical stage for learning the affective associations attendant to the development of motives and motivational expectancies. If the environment supports learning these associations during the critical period, then strong motivational dispositions will develop; if the environment supports a time earlier than the critical time, anxiety connected with these motivations will develop and if the environment supports learning at a date later than the critical time, very specific motives and expectancies will ensue. '
>
> (op. cit., p. 17.)

This emphasizes the specific contribution of the environment. Furthermore, it implies that, in research, we should treat the environment as a whole which has a specific meaning for the developing personality and intellect deriving from the specific configurations of the whole matrix of factors. This is an important point which may help to explain the inconsistencies of research findings. The usual psychological approach has been to isolate particular factors (specific techniques in punishment for example) and relate these to motivation or achievement. It was not understood that this specific research act carried an implicit assumption that the same techniques carried out in this way meant the same thing to the child regardless of the social context. It was thought of as a single ' stimulus ' similar to the stimulus which is produced in a laboratory by manipulation of an independent variable. Unfortunately, (or fortunately) family life is not a laboratory and this assumption cannot be made at present.* The ' situation ' must be seen as a matrix of interrelating factors which is destroyed when a single crude factor is abstracted since the factor derives its meaning from its context. For example, ' but you *must* go to school ', does not mean the same in a context of love and mutual respect as it does in a context in which the child fully understands the implicit relief of the parent in being rid of him. Yet the degree of firmness might be exactly the same.

---

*that middle class parents are more responsive to mass media child-rearing instruction.

Because of this problem in dealing with the environment†, it is to be expected that those studies which looked for a generalized picture of the family environment (' loving ', ' supportive ', etc.) and those which employed social class analysis, came closest to general agreement.

Before dealing with social class analysis we might mention a fairly elementary relationship which does appear to hold in all situations. Parental encouragement does have an important bearing upon school motivation and performance. Mannino (1962) reported that, with socio-economic status controlled, ' staying on ' at school after the minimum leaving age related positively to parental encouragement. Wall (1962) concluded that ' children, particularly boys, do very much better at school if their parents are interested in their progress ', and that ' children whose parents are rated as " very interested " improve their test-scores between 8 and 11 but those whose parents are " uninterested " show a deterioration in score '. Douglas (1964, p. 61) also suggested that children tend to work well when their parents take an interest in their school progress and to work badly when they are uninterested.

Social class values and achievement motivation

Many studies have shown a positive association between the level of educational and occupational aspirations on the one hand and various measures of social status on the other. Usually it is assumed that this association is a ' real ' one which results from the influence of the particular constellations of occupational, educational and action values which are implicit in the culture of the middle class. However, the equally clear relationship between school ability and middle class status is not usually accepted at face value. At least part of this association is thought to be due to an intervening variable called ' intelligence '. Sewell (1957) considered the possibility that the aspiration-social class relationship was also due to the intervention of this factor, intelligence. Tests on a random sample of high school seniors covering the state of Wisconsin confirmed that the relationship between aspiration and social class was not simply due to the common relationship of these variables to intelligence. Social class (whatever that is) makes an independent contribution. This is important because the conclusion was arrived at by controlling for ' intelligence ' which itself can be shown to be derived, to some extent, from sociocultural learning experiences. Thus, the importance of social status showed through even though it was weakened by controlling intelligence.

---

*This is not to say that the logic of research is wrong. The ideal aim is still to discover in the environment all the most important stimuli and to refine the techniques by which they can be measured. It only means that our actual abilities to do this in family life research covering a long period of time are incapable of handling the environment in this way at the present time.

†As a sociologist the writer must insist upon including the individual's perception of the social reality as part of ' the environment ' since psychologists often do their research studies serious damage by assuming that ' the environment ' constituted only the material and geographical aspects of the environment.

With Sewell as justification and Davis as inspiration we can now consider the value-systems of the different social classes to be learning environments which participate in the development of motivation and achievement levels. Davis had urged the use of class analysis in investigating the environment for factors in intellectual and personality development because it appears to provide the best means by which the cultural experience of the child may be described. It represents the best shorthand method for summarizing the individual's cultural beliefs, his definitions of life problems and possible solutions, the meaning of his language, and his culturally learned conceptions of the teacher, himself and school. As Kohn (1963, p. 471) puts it: ' It is so useful because it captures the reality that the intricate interplay of all these variables creates different basic conditions of life at different levels of the social order '.

Early Leaving was taking a similar, if much cruder, line when, in partial explanation of the fact that working class children leave grammar school before their middle class peers, it has the following to say:

> ' One of the main influences must be sought in the outlook and assumptions of parents in various walks of life. Consider first the outlook of parents in professional occupations. Most of them themselves received a grammar school or similar education, and others have made their way into a position in society in which they find such a background taken for granted. They are all engaged on work for which a fairly high level of education is an obvious advantage and many follow professions to which a specified educational standard is a condition of entry. In the circumstances it is not surprising that they assume that their children will not leave school at 15 but will stay as much longer as their ability justifies. This assumption is not due to any conscious sense of the value of education; it may be a mere social convention which has never been questioned. But it is in any case a powerful influence on parents towards keeping their children at school, and on children towards staying '.
>
> (op. cit., p. 36.)

Before *Early Leaving* was published, Kahl (1953) documented a similar phenomenon in the American educational system. During the course of a nation-wide mobility study it became plain that ' I.Q. and family status were useful predictors of the educational and occupational ambitions of high school boys '. However, within the lower-middle range it was found that these relationships were weaker than at the higher and lower reaches of the scale. Kahl, therefore, chose a pair of matched samples of boys from this status group, one representing boys who wished to go to college and another representing boys who did not wish to go. Interviews revealed that, ' although there was a common way of life which identified the common-man class, some members were content with that way of life while others were not '. Discontented parents, seeing that the educational system was a way out for their sons, tended to train them from the earliest years of grammar school to take school seriously. ' Only sons who internalized such values were sufficiently motivated to overcome the obstacles which faced the common-man boys in school; only they saw a good reason for good school performance and college aspirations ' (Kahl, 1953; p. 203).

An important finding of the study was that boys, to an extent not previously imagined, took up and held as their own, their parents' perceptions of possible opportunities, the desirability and possibility of status mobility, methods of obtaining mobility, and the goals which they felt appropriate to boys in their position at school. Neither the school nor mass media seemed to have much influence in moulding these ideas. This finding, which has found support in other research, is a very firm invitation to researchers to concentrate upon family environment, and since that time a great deal of research has been carried out on parental values and their effect upon motivation in school.

A further extension of this work within the Harvard Mobility Survey was undertaken by Elizabeth Cohen (1959) who studied fifty matched pairs of boys in the ' common-man ' group. Each pair consisted of a ' mobile ' and a ' non-mobile ' boy matched for I.Q., ethnicity, town of origin and parental membership of the common-man group. Attitudes of parents and sons were obtained by Guttman scale testing, and the basic aim was to chart the channels of parental influence upon the motivation of their sons. Cohen's theoretical model included three principal channels of influence; through socio-economic characteristics, through generalized attitudes towards college, and through concrete encouragement. In these ways the parents were expected to influence the decisions of their sons as to whether they wished to go to college or not.

In analysing the influence of socio-economic status, Cohen suggested that a significant difference in motivation is found among children of working class parents who have close ' on the job ' associations with people of higher status. As regards ' pressure ' it was found that parents of mobile sons seemed deliberately to encourage them to aim for middle class occupations and educational qualifications but did not push for higher school performance as such.

A similar sort of study was reported by Reissman (1953), aimed at determining the relationship between class and aspirations. Apart from age, he found that the two important aspects of the individual's value-system were his reference group and certain ' life-orientations '. The reference group is obviously important to the degree of satisfaction the individual has with past mobility and future aims. Thus a low level of aspiration can be based on a feeling of achievement in relation to the normal achievement of his reference group. Orientations other than success and mobility can be dominant for some individuals and thereby lead them to exert their energies in a different direction, or prevent them from aiming for mobility.

In England, Himmelweit (1955) enquired into the relationship between socio-economic background and behaviour on the one hand, and the formation of attitudes and values on the other. It is interesting to note here the use to which a social psychologist has put the Davis hypothesis. For Dr. Himmelweit and many American social psychologists, this hypothesis has suggested that ' a kind of adaptive socialized anxiety is generated more rapidly as a result of middle-class methods of upbringing '. The difficulty experienced by working class children to adapt to the middle class school culture was also considered. Information was collected on a wide variety of topics with the aim of ascertaining the degree to which Davis' hypothesis could be confirmed in England on representative samples of grammar school and secondary modern school children from middle-middle to lower working class.

The results confirmed some of Davis' and Warner's hypotheses. Middle class children were more concerned with how well they do at school, as are their parents. Parents also supervise the leisure of their children more. However—

> ' despite so many confirmatory results as to the greater pressure to which the middle class child is subjected, both externally by the parents and internally by his more rigid super-ego and more intra-punitive technique of handling frustrations, no evidence of greater overall anxiety or tension was obtained '. (Himmelweit, 1955, p. 34.)

A further set of findings was concerned with the transmission of values. One of the hypotheses was that the upwardly mobile working class boy would try to become like his middle class co-pupil with the result that his answers might be different from those of the non-mobile members of his class. It was also expected that the two middle class groups (successful and unsuccessful) would be similar. There were two main results. Firstly, the difference between the school groups of the middle and working classes were so much greater than those between social classes within the same school that differences were thought to reflect, in large measure, differences in educational and intellectual achievement. Secondly, the difference between the working class groups was most marked. Parents of successful working class boys had a much more middle class set of values than even the middle class.

Hyman (1953) saw a system of beliefs and values peculiar to the lower class acting as an ' intervening variable mediating the relationship between low position and lack of upward mobility '. The components of this value-system were:

1. less emphasis on traditional high success goals;
2. increased awareness of lack of opportunity to achieve success;
3. less emphasis on the achievement of goals which in turn would be instrumental for success.

These all add up to a self-imposed barrier to mobility. Thus, apart from the structurally imposed barriers to mobility they have a further set of hindrances in their own ideas about mobility.

Hyman analysed census and other data looking for evidence that motivation to advance to high places and to obtain the right sort of training for doing so was lower in the lower class. Considering first the value placed on formal education, he quoted four studies which supported the idea that the lower class always placed a lower value on education than those about them in the social scale. He found that there were crucial differences in values as regards motivation to advance in the economic structure. For example, when respondents were asked to choose their ideal occupation there was a strong tendency for upper class people to choose occupations for ' congeniality to person ' reasons. On the other hand, the choices of lower class people indicated much more interest in the economic benefits of the occupation. Only a very few wealthy people chose skilled occupations as being preferable to others, whereas a major proportion of lower class people chose this form of qualification. In these ways Hyman emphasized that there is reduced striving among the lower classes, an awareness of lack of opportunity, and different ideas as to how social mobility might come about.

The special relevance of values in the family was clarified by Rosen (1956) who expressed dissatisfaction with current psychological and sociological explanations of social mobility. Hitherto, says Rosen, the greater achievement of the middle class has been explained either by a process of social selection—the middle class being more intelligent, healthier and more attractive; or by the differential opportunities open to the two classes—the ' life chances ' hypothesis. These explanations, he agreed, were relevant and consistent with each other, but they were not exhaustive:

> ' since neither explanation takes into account the possibility that there may be psychological and cultural factors which affect social mobility by influencing the individual's willingness to develop and exploit his talent, intelligence and opportunities '. (Rosen, 1956, p. 203.)

Therefore, his thesis was that differential rates of mobility may also be explained as a function of differences in the motives and values of the social classes. Rosen saw that there were two components in the achievement orientations of the different classes. Firstly, a psychological factor—the personality characteristic of ' achievement motivation ', which provides an internal impetus to excel. Secondly, there is also a cultural factor consisting of certain value orientations which define and implement achievement motivated behaviour. He suggested that values affect mobility in three ways:

1. they provide a definition of goals;
2. they focus the attention of the individual on achievement;
3. they prepare him to translate motive into action.

Thus, whether or not the individual tries for mobility will be decided by his values. Before the drive for achievement can be expressed in the right, culturally defined terms there must be awareness of, and willingness to undertake, the necessary steps for achievement. These will of necessity involve the ability to plan, to sacrifice, and to strive consistently. Here again, abilities like these will be partly dependent on the values of the individual.

Rosen goes further to make a crucial distinction between widespread values and implementary values. Frequently we can see no difference between classes in the holding of some popular widespread value like the desirability of upward mobility. But a difference is found in the implementary values which facilitate and encourage achievement. He summarized the situation in America in terms of basic or implementary values in this way:

> ' Middle class children are more likely . . . to embrace the achievement value-system which states that given the willingness to work hard, plan and make proper sacrifices, an individual child should be able to manipulate his environment so as to ensure eventual success '. (Rosen, 1956, p. 211.)

Progress, that is, is not necessarily a question of drive. At least as important is the way in which the individual fits in with the requirements of his present situation. This is obviously so with our education system. Successful mobility can come to those who simply adapt themselves well to each successive stage, the system itself carrying them up as on the crest of a wave. This meshing in with the values, perceptions and behaviours of a culture-pattern constitutes, with the achievement drive, the two basic factors which produce mobility within the framework laid down by the structure of possibilities.

Having established the abundance of theoretical justifications for an analysis of family value-patterns along social class lines, it now remains for us to deal with some of the more recent substantive findings of research studies which have been designed on these lines.

Douvan (1956) sets the stage with an experiment designed to test the hypothesis that 'achievement want' of middle class children is more generalized than that of working class children and hence is less likely to vary with changes in the reward offered for achievement. Her conclusion was that the achievement want of the middle class child was more generalized as well as more dependable.

Douvan (1958) also provides us with a useful thumbnail sketch of the upward and downward-aspiring boy. In this survey, the authors classified boys as upward mobile, stable, or downward mobile, depending upon whether their occupational aspirations were higher than, similar to, or lower than their father's occupation. Interviews with the boys revealed the following categorizations. The upward-aspiring boy was described as being unusually lively and energetic. He expressed concern for the intrinsic interest of his future work and has a longer time-perspective when thinking of occupations. A set of pictures was used to discover the degree and form of internalized standards possessed by the two extreme groups. These pictures portrayed an adolescent boy facing conflicting pressures. For example, he might be torn between a promise to parents to be home at a particular hour and a wish to accept pressure from his friends to stay out late. In both groups two-thirds of the respondents decided that the boy would go home. Thus there was no difference in prognosticated *behaviour*. However, there was a difference in perceived motives for the action in that the upward-aspiring boys tended to explain his actions in terms of his promise to his parents, or because his parents had trusted him. The downward-aspirers were more likely to explain his return by fear of punishment. The implications of these findings for the authors was that the upward-oriented have internalized personal controls in contrast to the downward-oriented who showed a tendency to internalize external standards and rebel against them. The upward-aspirer was found to show greater independence from his family whereas the downward-aspirer was at once more dependent on, and rebellious towards, his family. Finally, the upward-aspirer showed high self-acceptance and social confidence in comparison with the downward-aspirer who was more ambivalent towards himself and more unsure in social situations.

These findings fit neatly with those of Kohn (1959) who showed clearly that child-rearing values of middle class parents stress internalized standards for governing relationships with other people while the working class concentrated upon qualities of honesty and neatness. Apart from demonstrating the importance of social class analysis in this question (as does Kohn, 1960-1961) the findings support Duvall's description of working and lower middle class parental values as 'traditional' in comparison with the middle class 'developmental' emphasis. With the working class parent it is the actual act that matters. The child must accept externally imposed rules. To the middle class parent it is the motives of the child which give meaning to the act. Thus we can expect that where the middle class family is successful in transmitting such a point of view to the child, he will develop a responsibility for his own values and actions which would support his acceptance of the parental views on educational and occupational aspirations. On the other hand, a working class point of view on education (most social classes feel that education is A GOOD THING) will tend to be imposed upon the child together with other external rules. There is greater likelihood that the child will rebel against this as against other arbitrarily imposed rules. There is then a further factor that even where the working class child internalizes the value it will mean something different to him. Horizons will be lower and understanding of the situation and its requirements less efficient.

Katz (1964) studied the meaning of success among Australian adolescents and found evidence of considerable variation according to socio-economic class of origin. These differences appeared to indicate ' a major difference in the frame of aspirational reference internalized by members of different social classes '. In the case of the middle class adolescent, the success goal was prestige achievement attained through personal effort and worthiness. Adolescents from the unskilled working class, on the other hand, had little concern for social status. Success is defined in terms of ' possessions which are procured by personal exertion, but limited in possibility of attainment by factors over which the individual has no control '. Between these two extremes, adolescents of skilled worker families showed a great deal of variation, suggesting to Katz that some had internalized middle class and others, unskilled class, perceptions.

The importance of chance in the lower working class perception of the environment is supported by research in Britain and most relevantly by the theories and research of Basil Bernstein (1962). On the basis of his analysis of linguistic forms, Bernstein suggests that middle class life and language encourage the development of an ' instrumental ' attitude to the world around which is distinct from the more intellectually passive perceptions of the working class.

We have seen that the values of working class life tend to demand adoption of rules which are seen by the child to be external and arbitrary (' because I say so '). We also know that the horizons of social actuality implicit in the specific behavioural imperatives by which working class parents seek to organize the behaviour of children are ' shorter ' and ' narrower ' than those of the middle class families [though none the less ' accurate ' for that (Liversidge, 1962)]. Since middle class children have higher need achievement in relation to education and this clearly is important in actual performance*, we can assume that the imposition of arbitrary rules and shorter horizons are both depressive factors in the need for educational achievement. They clearly are when taken together and in certain cultural situations. However, they need not go together. In line with the earlier theoretical objections to abstraction of a single factor for treatment as a continuously operating variable, we have to consider the possibility that different situations will give very different functions to any single factor.

The lower middle class provides a useful test of this possibility. Duvall (1948) had joined the lower middle class to the working class in categorizing their child-rearing ethos as ‘traditional’. On the other hand, the work situation, upon which one might expect their achievement values to depend, would be different. They would tend to do work which demanded skills in the manipulation of ideas and persons as opposed to working class activities with things. Promotion will, superficially at any rate, appear to depend upon the actual performance of the individual. Finally, the lower middle class is likely to be less subject to control than the working class.

---

*Douglas (1964, p. 60) found that according to teachers, 26% of upper middle class children work ' very hard ' as opposed to 7% of upper and lower working class: (p. 64) children who work ' hard ' or ' very hard ' get 12% more grammar school places than they would be expected on their tests to get, while ' poor ' or ' average ' workers got 25% fewer places than their tests would lead us to expect.

With all these points in mind, and in the light of Argyle's (1962) suggestion that introjection or projection of parental ambition is important in parent-child relationships, Swift (1964) related the 11+ performance of lower middle class children to the mobility pessimism of their fathers. In a ten per cent random sample of a Local Education Authority Division it was found that, within the middle class, the father's dissatisfaction with his job and its prospects related significantly to the likelihood of his child's success on the 11+. In addition it was found that mobility-pessimism was very dependent upon lower middle class membership. In the sample only one out of nine lower middle class fathers of successful children did not have ‘high’ mobility-pessimism.

On the basis of these very clear but necessarily tentative findings, a second ideal-type of academically successful family was constructed.* The principal intention behind this was to emphasize the weakness of the position which tends to see ‘the family environment’ as a variable which ranges from ‘good’ to ‘bad’ in terms of its capacity to produce children who are adaptive to the needs of the educational system. This is a logically reasonable perspective to hold but it carries with it the danger of implicit assumptions about the linearity of relationships between its component elements and educational adaptiveness which are not necessarily valid. In this case, research has emphasized the importance of a supportive, ‘democratic’ family life which encourages the development of personal confidence and satisfaction with values which are accepted rather than imposed. The assumption which tends to follow from this is that the ‘developmental’ environment is best at the top and non-existent at the bottom of the social class hierarchy. In this way it can be seen to be a stimulus which relates in a linear fashion to need for educational achievement and hence educational success.

This second ideal-type of academically successful family was one in which mobility-pessimism was high, and discipline was ‘traditional’. Commitment to education was ‘high’ but of a very different nature to the high commitment expressed in the ideal-type middle class family where it is valued for its liberating qualities for the individual. Instead (in the light of the ‘traditional’ discipline ethos) it was seen as a necessary aspect of certificate-collecting which was, in turn, seen to be the major ‘external’ demand of the social environment. One way in which this family differs from the working class family lies in the social horizons which the parents hold for their children. Because of similarities in work situation and content, and probably of early educational experience, the lower middle class parent will have horizons which are closer to those found in the remainder of the middle class. He is also likely to understand what is involved in school work and will not be burdened by the idea that the people who get into the higher classes represent a different sort of person from himself.*

---

*An alternative, that is, to the usual ideal-type of ‘cultured’ upper-middle class family in which the child is stimulated in a generally tolerant fashion to develop his capabilities in many directions.

This all adds up to a very different sort of environment to the stereotypical ' educationally successful ' middle class family. It is also one which cannot logically be placed mid-way between skilled working and middle class except on a single specific criterion like income. According to the hypothesis, the nearer a family approaches this ideal-type, the more likely is their child to be successful in the school. The whole analytical situation is then made vastly more complicated when we remember that the hypothesis has been framed in terms of two ideal-types, neither of which is likely to exist. Any given family will simultaneously go some way towards both ideal-types. However, social class analysis can help in ' controlling ' for certain configurations of variables to the extent that work-situation, income and education tend to encourage their development.

Conclusion

An attempt has been made to present a perspective on the relationship between family environment and the need for educational achievement. This is a sociological perspective which employs the concept of values as the link between the social environment on the one hand and educational motivation on the other. It is suggested that research upon the development of educational potential has tended to oversimplify the subtle relationships between the developing personality and its environment.

However, research and theory which recognized the coherence of the cultural environment have suggested some general propositions. The accuracy of these propositions may not prove to be high, but if British research follows the lead given by American workers, away from descriptive and naïve correlational studies towards socio-cultural analysis, there is real prospect of light.

Bibliography

Argyle, M. and Robinson, J. (1962). ' Two Origins of Achievement Motivation ', *Brit. Jour. Soc. and Clin. Psych.*, vol. 1, part 2.

Bene, Eve (1959). ' Some Differences between Middle Class and Working Class Grammar School Boys in their Attitudes towards Education ', *Brit. Jour. Sociol.*, vol. X.

Bernstein, Basil (1962). ' Social Class Linguistic Codes and Grammatical Elements ', *Language and Speech*, vol. V, pp. 221-240.

Bronfenbrenner, Urie (1958). ' Socialisation and Social Class through Time and Space '. *In:* Maccoby, E. E., Newcomb, T. M. and Hartley, R. L., eds., *Readings in Social Psychology*, pp. 400-425. New York: Henry Holt.

Campbell, W. J. (1955). ' The influence of Socio-Cultural Environment on the Progress of Children at the Secondary School Level ', *Australian J. Psych*, vol. VII, pp. 140-146.

Centers, Richard (1948). ' Motivational Aspects of Occupational Stratification ', *J. Soc. Psych.*, vol. 28, pp. 187-214.

Clausen, J. A. and Williams, J. (1963). ' Sociological Correlates of Child Behaviour '. *In:* Stevenson, H. W., Kegan, J. and Spiker, C., eds., Child Psychology section in *Yearbook of National Society for Studies in Education*, vol. 62, pp. 62-107.

Cohen, E. (1959). ' Parental Factors in Educational Mobility '. Unpublished Ph.D. thesis, Harvard University.

Cohen, E. (1965). ' Parental Factors in Educational Mobility '. *Sociol. of Educ.*, vol. 38, no. 5, pp. 404-425.

---

*For the reasons which Hyman (1953) made clear, this has a depressing effect upon aspiration and action.

Davis, Allison (1948). *Social Class Influences Upon Learning*, Cambridge, Mass.: Harvard University Press.

Davis, Allison (1949). 'American Status Systems and the Socialisation of the Child'. *In:* Kluckhohn, C. and Murray, H. A., *Personaltiy in Nature, Society and Culture*. New York: Alfred Knopf.

Davis, Allison (1960). *Psychology of the Child in the Middle Class*. Pittsburgh: University of Pittsburgh Press.

Douglas, J. W. B. (1964). *The Home and the School*, pp. 60-65 and 159-162. London: MacGibbon and Kee.

Douvan, Elizabeth (1956). 'Social Status and Success Strivings', *J. Abnorm. and Soc. Psych.*, vol. 52, pp. 219-223.

Douvan, Elizabeth and Adelson, J. (1958). 'The Psychodynamics of Social Mobility in Adolescent Boys', *J. Abnorm. and Soc. Psych.*, vol. 56, pp. 31-44.

Drews, Elizabeth Monroe and Teaman, John E. (1957). 'Parental Attitudes and Academic Achievement', *J. Clin. Psych.*, vol. 13, pp. 328-332.

Duvall, Evelyn M. (1946). 'Conceptions of Parenthood', *Amer. J. Sociol.*, vol. LII, pp. 193-203.

Elder, Glen H. (1965). 'Family Structure and Educational Attainment; A Cross-National Analysis', *Amer. Sociol. Rev.*, vol. 30, no. 1, pp. 81-96.

Floud, J. E., Halsey, A. H., and Martin, A. M. (1956). *Social Class and Educational Opportunities*. London: Heinemann.

Fraser, Elizabeth (1959). *Home Environment and the School*. London: University of London Press.

Himmelweit, Hilde T. (1955). 'The Psychological Aspects of Social Differentiation', *Inter. Soc. Science Bull.*, vol. 7,

Hyman, Herbert H. (1953). 'The Value Systems of Different Classes'. *In: Class Status and Power*, Bendix, R. and Lipset, S., eds., pp. 426-442.

Kahl, J. A. (1953). 'Educational and Occupational Aspirations of "Common Man" Boys', *Harvard Educ. Rev.*, vol. 23, pp. 186-203.

Katz, F. M. (1964). 'The Meaning of Success: Some Differences in Value Systems of Social Classes', *J. Soc. Psych.*, vol. 62, pp. 141-148.

Kohn, Melvin (1959). 'Social Class and Parental Values', *Amer. J. Sociol.*, vol. LXIV, pp. 337-351.

Kohn, Melvin (1963). 'Social Class and Parent-Child Relationships: An Interpretation', *Amer. J. Sociol.*, vol. 69, pp. 471-480.

Kohn, Melvin L., and Carroll, Eleanor E. (1960). 'Social Class and the Allocation of Parental Responsibilities', *Sociometry*, vols. 23-24, pp. 372-392.

Liversidge, William (1962). 'Life Chances', *Sociol. Rev.*, vol. 10, no. 1, pp. 17-34.

Mannino, Fortune V. (1962). 'Family Factors Related to School Persistence', *J. Educ. Sociol.*, vol. 35, no. 5, pp. 193-202.

Ministry of Education (1954). *Early Leaving*. London: H.M.S.O.

Reissman, L. (1953). 'Levels of Aspiration and Social Class', *Amer. Sociol. Rev.*, vol. 18, June, pp. 233-242.

Rosen, Bernard C. (1956). 'The Achievement Syndrome: A Psycho-cultural Dimension of Social Stratification', *Amer. Sociol. Rev.*, vol. 21, April, pp. 203-211.

Rosen, B. C., and D'Andrade, Roy (1959). 'The Psycho-social Origins of Achievement Motivation', *Sociometry*, vol. 22, pp. 185-218.

Sewell, William H., Haller, Archie O. and Murray, A. (1957). 'Social Status and Educational and Occupational Aspiration', *Amer. Sociol. Rev.*, vol. 22, pp. 67-73.

Swift, D. F. (1964). 'Who Passes the 11+?', *New Society*, 5th Mar., pp. 6-9.

Swift, D. F. (1965). 'Meritocratic and Social Class Selection at Age Eleven', *Educ. Res.*, vol. VIII, no. 1, pp. 65-73.

Toby, Jackson (1957). 'Orientation to Education as a Factor in School Maladjustment of Lower Class Children', *Social Forces*, vol. 35, pp. 259-266.

Veroff, Joseph (1965). 'Theoretical Background for Studying the Origins of Human Motivation Dispositions', *Merrill-Palmer Quarterly of Behaviour and Development*, vol. II, no. 1.

Wall, W. D. and Miller, K. M. (1962). 'Motivation and Counter-Motivation'. *In: Child and Education*, Proceedings of XIV International Congress of Applied Psychology, vol. 3, pp. 161-175.

EXAMINATIONS IN RELIGIOUS EDUCATION–STATISTICS AND INFERENCES

by J. G. Harris

Department of Religious Education,
Monmouthshire College of Education

A GOOD deal of recent educational thought and practice has been concerned with ways and means of stimulating progress and measuring attainment over the whole curriculum field. The ideal method of doing this has not yet been evolved, but the ease with which we have succumbed to the lure of examinations, and the importance they assume in our educational system, is obvious once we glance at the numbers of pupils we subject to the rigours of the examination processes. In 1950 we examined 809,457 pupils for the School Certificate, whereas ten years later the figures had soared to 1,603,894 in the G.C.E. 'O' level examinations. In 1950, 101,733 sat the Higher School Certificate examination, but in 1960 the number was 213,993. These figures are for the summer examinations of the G.C.E. for the whole country, but regional boards, like the Joint Matriculation Board, report a similar increase. Whereas this board examined 986 candidates for the Higher School Certificate in 1920, in 1930 the figure was 3,480, in 1950, 12,278, but in 1960 21,302, and in 1964 31,212 (the last two figures being for the 'A' level examinations).

In general, examinations in their present form are an accepted part of our educational system, and few educationists question their propriety, even though they may call for radical revision in certain directions. But with regard to Religious Education the situation is very different. There is a body of opinion which would eliminate examinations in religion altogether in order to concentrate on the nurture of religious experience. On the other hand there are teachers of religion who believe that examinations give the subject equal status with others and provide a stimulus to effort and a reward for progress.

The purpose of this paper is twofold: 1. to demonstrate by available statistics the growth of examinations in Religious Education, from which we may deduce an interest in the subject for its own sake, and the value of examination qualifications in it; 2. to indicate how this provides a basis for further experimentation in this field.

The increase in numbers of examination candidates

The following statistics of the national trends in the numbers of candidates sitting G.C.E. examinations in Religious Education at 'O' and 'A' level show a notable increase since 1951:—

1951 : 22,497 entered 'O' level examinations in R.E. (9,324 boys, 13,173 girls) out of a total entry in all subjects of 738,717
557 entered 'A' level examinations in R.E. (148 boys, 409 girls) out of a total entry in all subjects of 103,803.

1955 : 27,664 entered ‘ O ’ level in R.E. (9,432 boys, 18,232 girls) out of a total entry in all subjects of 1,043,864.
871 entered ‘ A ’ level examinations in R.E. (256 boys, 615 girls) out of a total entry in all subjects of 137,867.

1960 : 44,097 entered ‘ O ’ level examinations in R.E. (14,683 boys, 29,414 girls) out of a total entry of 1,603,894.
1,984 entered ‘ A ’ level examinations in R.E. (580 boys, 1,404 girls) out of a total entry of 143,817.

1962 : 52,770 entered ‘ O ’ level examinations in R.E. (18,505 boys, 34,265 girls) out of a total entry of 1,832,949.
2,899 entered ‘ A ’ level examinations in R.E. (832 boys, 2,067 girls) out of a total entry of 269,710.

Between 1951 and 1962 the entry in Religious Education rose by about 140% in the ‘ O ’ level examinations, which is roughly proportionate with the increase in the total entry to this examination. Religious Education maintained its attraction as an examination subject at this level, and in certain cases attracted more candidates proportionately than any other subjects with which it may be legitimately compared for the purpose in hand. For example, history, which requires learning ability and the capacity to reproduce a body of factual information, attracted 66,250 candidates in 1951, in 1955 84,858, in 1960 118,671, and in 1962 134,086, an increase of about 100% between 1951 and 1962. Other subjects of a comparable nature show a similar trend.

We notice from the figures that the increases are greater amongst girls. But looking at the ratio of girls to boys who were at school at the age of sixteen, the normal age for taking the ‘ O ’ level examination, in relation to the national population, in 1951 the boys of this age still at school formed 41.2% and the girls 37.7%, but in 1962 the comparable figures were 76.2% boys and 65.3% girls, of the total population, so that it must be assumed that examinations in Religious Education have proved more attractive to boys than to girls at this age.

The national figures show a remarkable increase in the number of candidates taking ‘ A ’ level examinations in Religious Education, a notable fact when it is remembered that the sixth form is usually geared to intensive specialization. In fact the increase at this level is more striking than at the ‘ O ’ level. Five times as many candidates sat the examination in Religious Education in 1962 than in 1951, the increase in the numbers of boys and girls being roughly equal. The comparable figures in history for the same years are 10,230 in 1951 and 20,177 in 1962, approximately a 100% increase. In English literature 12,923 candidates entered for ‘ A ’ level in 1951 and 26,414 in 1962, again just over 100% increase. In neither case is the increase in candidates proportionately as great as in Religious Education. In relation to the total entry to all subjects at ‘ A ’ level, Religious Education attracted 0.54% in 1951 and 1.1% in 1962. Although Religious Education as an examination subject at sixth form level appeals only to a minority, this appeal has held its ground and increased during the past decade or so.

Levels of performance

Having established the main trends in the numbers entering for external examinations in Religious Education during the past ten or more years we may look at the level of performance at 'O' and 'A' level. This is a statistical survey only, and a full analysis would require detailed estimates of the examination syllabuses and the demands made upon the candidates in different subjects. It may be said, however, that the Religious Education examination syllabuses compare favourably with syllabuses in other comparable subjects in respect of content and the level of difficulty. In particular, the schemes of study and the standards of examinations in Religious Education compare favourably with subjects like history, English literature, which require learning ability and comprehension of meaning, and with mathematics, which requires a measure of inductive reasoning ability.

Turning to the national figures, the table of performance reads as follows:

1951 : entered at 'O' level 22,497; passed 52.5% (boys 39.7%, girls 61.5%)
entered at 'A' level 557; passed 78.5% (boys 66.9%, girls 82.6%)
1955 : entered at 'O' level 27,644; passed 56.6% (boys 43.2%, girls 63.5%)
entered at 'A' level 871; passed 80.5% (boys 71.5%, girls 84.2%)
1960 : entered at 'O' level 44,097; passed 55.8% (boys 44.1%, girls 61.7%)
entered at 'A' level 1,984; passed 65.4% (boys 55.0%, girls 69.7%)
1962 : entered at 'O' level 52,770; passed 58.4% (boys 48.7%, girls 63.6%)
entered at 'A' level 2,025; passed 69.9% (boys 64.3%, girls 72.1%)

No very significant improvement in the performance of candidates can be claimed during the years here being reviewed, although there is a steady upward trend. It is worth noting that whereas the attainment of girls has remained steady, although higher than that of boys, there is a marked appreciation from 39.7% in 1951 to 48.7% in 1962 in the successes registered among the latter. But the position is very similar in other comparable subjects. In 1951 57.2% were successful in history and 58.6% in 1962; in English literature in 1951, 59.6% passed, and in 1962, 58.2%; in mathematics the passes in 1951 were 57.0% and in 1962 57.2%.

The increase in the entry to the 'A' level examination in Religious Education between 1951 and 1962 is notable, from 557 to 2,025, but there has been a depreciation in the level of performance from 78.5% passes in 1951 to 65.4% in 1962. This is a serious decline, although it accords with the national tendency over the whole examination during these years. Some attribute the diminution of achievement at the 'A' level to the raising of standards or to the wide range of ability with which inflated sixth forms have to deal. How far the nature of the syllabus, the demand for critical appreciation, and the pressure of specialization account for this is a matter for further investigation. Or to what extent a new approach to sixth form teaching would contribute to higher examination performance in all subjects is a subject that needs to be constantly reviewed in the light of changing conditions. It will be noted that the performance of boys falls below that of girls, although the difference is not as great as at the 'O' level.

When we compare the performance in some other subjects at 'A' level with that in Religious Education we have the following figures. Whereas in 1951, 78.5% of the candidates passed in Religious Education, in history the figure was 72.1%, in English literature 76.3%, and in mathematics 73.3%. In 1955 the respective figures were Religious Education 80.5%, history 76.3%, English literature 70.3%, mathematics 68.7%; in 1960 Religious Education 65.4%, history 71.3%, English literature 70.3%, mathematics 68.7%; in 1962 Religious Education 69.9%, history 71.7%, English literature 71.5%, mathematics 69.3%.

Discussion

No one who upholds the propriety and value of examinations in Religious Education could sit complacently before the present figures. The level of failure is far too high, as indeed it is in other subjects, but this poses the question of how far the examinations in Religious Education as at present conceived meet the needs and abilities of the candidates themselves. This question is under active review and requires the serious consideration of teachers of religion.

The interest in examinations in Religious Education is well demonstrated by the statistics. Having regard to the nature of the subject, and the rightful aims of religion in school, what form of examination is best suited to the nature of so personal and individual a subject as this? Granted that religion is sustained by a solid body of historical facts, and that it is possible to examine a candidate's grasp of the factual and historical basis of religion, can we equally fairly and objectively examine his understanding of religious truth, the meaning and nature of religious experience, the relevance of religion to his existential situation, and the application of religious insights to the contemporary scene? If religion is examinable in a way that does not violate its essential nature, who is capable of conducting such an examination without introducing his own personal understanding and predispositions? Is an examination that requires only factual learning, and whose approach is entirely academic, the right kind of religious examination?

To raise these questions at this time is important if we are to pioneer a new approach to religious examinations. If it is accepted that Religious Education is about to enter upon a new phase requiring a more educationally congenial approach, and making greater use of the processes of learning and the nature of religious experience, considerable thought will need to be given to the schemes of examination to be followed. For example, what other examinable approach to the teaching of the Bible is possible than the traditional one of academic and historical study? How far can justice be done by the individual and corporate experiences of religion in any scheme of examination? If religion is to become an ' open ' subject in school, engaging pupils in dialogue with other school disciplines, and with the living issues of their own life and experience, how far can such development be represented in any examination scheme?

If, from the foregoing statistics we may deduce a real and slowly growing interest in religious examination, and we may assume that the subject appeals to the students' interest for its own sake, and not as a ' soft option ', the pioneering of the next stages will make heavier demands upon workers in this field. In particular, we shall need an examination system that will serve the aims of Religious Education, which is in line with contemporary thinking about the content and method of religious teaching, and which does not violate the contribution of religion to the full development of the child's personality.

SOME PROBLEMS OF EVALUATING i.t.a. A SECOND EXPERIMENT

by John Downing and Barbara Jones*

Reading Research Unit, University of London Institute of Education

Introduction

> '*Almost invariably in a scientific investigation, when one sets out to confirm or confute some plausible theory, the result is seldom the plain straightforward " Yes " or " No " that cross-examining council love to demand. Usually it appears that there are elements of truth in both the opposing views: and more often than not the most rewarding results are the discoveries made by the way. Fresh problems, quite unforeseen, nearly always emerge: new facts, quite unsuspected, are vividly brought to light.*'

Thus Burt (1962) wrote about the Initial Teaching Alphabet (i.t.a.) experiment when it began in 1961. Now, four years later, we at the Reading Research Unit are in the thick of the ' new problems ' which Burt predicted, and our first really detailed report on the original experiment is in the final stages of preparation.

Meanwhile a number of controversial articles have been written on the i.t.a. research, some favouring i.t.a., others against the new writing-system devised by Sir James Pitman (1961). Interim reports from the Reading Research Unit urged caution until more definitive reports could be published, and we have supported† others who have suggested that, pending these reports, educators should defer final judgement on i.t.a.'s effectiveness in our schools. Sometimes, however, writers who have suggested that teachers should be cautious have themselves displayed a lack of caution. This paper attempts to clear up misconceptions which may have been created by Southgate's (1965) recent article in this journal. It also includes a report of a second experiment with i.t.a. which we began in 1963, in an attempt to extend control over certain incidental variables including some of those to which Southgate referred.

Problems of evaluating i.t.a.

The title and purport of Southgate's article suggested that educators should be ' Approaching i.t.a. Results with Caution '—a view we endorse and have expressed repeatedly in our earlier publications on the i.t.a. research.‡ We agree with Gans (1964) that:

> ' The i.t.a. approach is at its beginning. As yet there are not enough results from the experimentation going on to warrant the extravagant claims that are being made for it. Only broad and longitudinal evaluation which yields reliable conclusions can determine its permanent place in the teaching of reading '.

---

*Special thanks are due to Stanley Halliwell, William Latham and Brian Lewis, all of the Reading Research Unit, for critically reading and commenting on an earlier draft of the paper. Many of the ideas contributed by these persons have been freely incorporated in the final version.

†For example, Downing, J. A. (1965). ' Current Misconceptions about i.t.a. ', *Elementary English*, May 1965, pp. 492-501.

‡For example, in the collection of interim reports (Downing, 1964) published last year, we said (p. 96): ' It must be emphasized that we do not consider that our research has reached a point at which *final* answers can be given . . . We would caution educators not as yet to form definite judgements or final conclusions regarding the use of i.t.a. '.

This ‘ broad and longitudinal evaluation ’ is being conducted by the Reading Research Unit, and the first definitive report of the large-scale experiment begun in September 1961 is currently being prepared for publication.

In 1968 it is planned to publish a similar report on the attainments of the same children after the transfer from i.t.a. to traditional orthography (t.o.). Until these reports are available no conclusions can be regarded as based on adequate empirical research findings. Even then there will undoubtedly be many unanswered questions.

As Southgate very properly points out, ‘ sober appraisal of what the results of this project may mean ’ is very difficult when a detailed account of the experimental design and the actual procedure followed by the experimenters has not yet been published. Such an account of the design and method of the main i.t.a. experiment (not only what was attempted but what actually happened) will form a part of the report to be published in 1966. It is probably the absence of such a full report which has led Southgate to make several incomplete comments on the design and execution of our experiment.

Before discussing these, however, we should examine the unusual conceptual framework used by Southgate in her discussion of the various motivating forces which she believes may have influenced the results of the main experiment begun with i.t.a. in 1961. As well as the ‘ Hawthorne Effect ’ Southgate suggests that we should consider other possible sources of bias in reading experiments—including the ‘ Reading Drive ’, the effects of visitors to the experimental classes, and publicity for i.t.a.

Southgate’s concept of a ‘ reading drive ’ is somewhat idiosyncratic: ‘ the main factor [accounting for results in the experimental classes] is the effects of what might be termed a ‘ drive ’ in the teaching of reading . . . A reading drive is an extremely potent force ’. A ‘ reading drive ’ is described in Southgate’s article as follows:

> ‘ A reading drive is basically a new surge of inspiration through the teaching of a subject. It ferments in the teacher and bubbles over on to the children who are thrust forward on its waves of enthusiasm. The drive is accompanied by an increase in interest, motivation and application which is inevitably followed by an improvement in attainments. A reading drive can stem from one ingredient or from a combination of many ingredients. These may include a new teacher or head teacher, a new scheme or method, new books, new apparatus, a new library, testing, recording, regrouping or cross classification of children, lectures and discussion groups for teachers ’.

This is a popular use of ‘ drive ’ and differs from the use of the word in its strictly technical psychological sense—e.g. a hunger drive.

Even if this concept of reading drive were refined to permit its consideration as a possible ‘ intervening variable ’ in reading research, a great deal of empirical work would need to be done to determine the relative potency of the various ingredients listed by Southgate. But instead of using a vague umbrella term it would perhaps be more useful to differentiate the various motivational factors present in such teaching situations.

In her discussion of reading drive Southgate comments on certain aspects of our original i.t.a. experiment. These comments may prejudice rather than aid judgement since they are based upon an inaccurate picture of what was actually done in the course of our experiment. For example, Southgate claims that 'The provision of equipment differs for experimental and control classes', and suggests that this difference entirely favoured the experimental group: 'The experimental classes have been supplied with complete new *Janet and John* schemes, including supplementary apparatus, all printed in i.t.a. They have also received stocks of new books for the book corners'.

The truth about the provision of books and other materials to experimental and control classes in our main experiment is more complex than this. The 'Field Experiment'[*] approach which has been followed in our first main experiment requires the maintenance of everyday-life conditions as far as is possible within the limits imposed by the need for control of variables. In order to keep close to the real-life classroom situation of Britain it was decided to use the *Janet and John* series—the most popular series of basic readers in this country. The experimental classes had their *Janet and John* books printed in i.t.a. while the control classes continued to use *Janet and John* books in traditional orthography (t.o.). In addition to the basic readers the various items of teaching apparatus published with this series were also produced in i.t.a. for the experimental classes.[†]

In British infants' schools much emphasis is placed on the importance of the Book Corner or classroom library, and it was recognized from the outset that, if the experimental conditions were to be representative of the normal approach in this respect, it would be necessary to produce many books in i.t.a. for the Book Corner. It must be realized that the matching of the conditions in respect of materials fell short of the ideal in our *initial* experiment. The children in the i.t.a. classes were at a *dis*advantage in respect of the shortage of supply of i.t.a. materials, and there were delays in supplying the schools with properly produced i.t.a. books for the library corner. To make good this shortage the Reading Research Unit duplicated sheets of i.t.a. text which the teachers pasted over the original t.o. printed pages. The effect of this 'pasting-in' of i.t.a. pages into books which were often already fatigued can be imagined. On the other hand, the i.t.a. *Janet and John* books were new and fresh in the experimental classes, whereas in only some of the control classes were the t.o. *Janet and John* books new and fresh. Thus many control schools continued working with used copies of the t.o. *Janet and John* books. It is possible, to some extent at least, that these differences have balanced out. The i.t.a. classes had new and fresh basic readers but suffered from a restricted choice of Book Corner materials, many of which were of the improvized 'paste-in' variety. The t.o. classes often had 'tired' copies of the basal readers but had a much wider selection of materials for the Book Corner.

In accordance with our aim to make a 'Field Experiment' approach, we considered other aspects of the child's environment in respect of books. In order not to disturb normal conditions in the experimental group it was decided not to attempt to isolate children learning i.t.a. from their normal t.o. environment. If the teacher in the infants' classroom normally used reference books which could not be produced in i.t.a. because of the cost, these remained in the classroom in their normal t.o. edition. Also, of course, children in the i.t.a. classes continued to meet t.o. at home and in the streets. From the point of view of the early learning stage the i.t.a. children were thus at a disadvantage because they did not have their classroom experiences reinforced to the same extent by experiences with printed English outside school. To offset this in some degree, i.t.a. materials were made available outside school for the experimental group in two ways. Public libraries in the locality kept a stock of i.t.a. books and head-teachers arranged for parents to buy i.t.a. books for their children's birthdays and for Christmas.

---

[*] The 'Field Experiment' developed from Lewin's (1951) proposal that 'we shall have also to develop research techniques that will permit us to do real experiments within existing "natural" social groups'.

[†] We wish to acknowledge the generous way in which Mr. Mackenzie-Wood, Chairman of the publishers, Messrs. James Nisbet & Co., co-operated with us in this respect.

Essentially, then, the difference between the experimental group and the control group was not one of a pure contrast between the former using i.t.a. materials and the latter using t.o. materials. Actually the situation was that the experimental group had experience of both i.t.a. and t.o. while the control group's experience was limited to t.o.

Probably the experimental group's experience of both i.t.a. and t.o. may have had an important effect on these children's ability to transfer from i.t.a. to t.o. at the later stage, but in judging the results of the earlier phase when the experimental group was using the i.t.a. basic series and being tested on i.t.a. tests, it must be taken into account that in the first two years of our initial experiment the i.t.a. children seem likely to have been at a disadvantage in terms of frequency of exposure to the letter and word forms they were learning. By the third year (1963), when over a hundred experimental i.t.a. schools had entered the research, the i.t.a. book situation had improved. Now, in 1965, a much greater variety of books is available in i.t.a. and it would be possible to provide a more adequate i.t.a. book environment for an experimental group (see, for example, the National Book League's* list of i.t.a. books). However, in evaluating our main i.t.a. experiment it must be remembered that supplies of i.t.a. books in 1961 and 1962 were very limited and improvization of the kind described above was necessary. Certainly it is misleading to suggest that the i.t.a. classes had a definite advantage in respect of the quality or quantity of books available.

A second issue raised in Southgate's article is the degree of success achieved in our unusual attempt to control the Hawthorne Effect in our main experiment. She states that 'The experimental group teachers attend regular monthly meetings to consider problems and discuss progress', but 'Few refresher courses have been provided for teachers in control schools nor have "regular meetings" been usual'.

Since i.t.a. was a highly novel approach, some features of which were a source of considerable misgiving to teachers and parents as well as the experimenters in 1961, we felt bound to keep a close watch on developments to safeguard the interests of the children. We tried, therefore, to hold meetings of the teachers of the i.t.a. experimental classes as often as possible at first in order to bring out in discussion any difficulties which they were experiencing. These difficulties turned out to be chiefly concerned with such matters as delays in supplying i.t.a. materials. These meetings of the experimental group teachers were conducted by a member of the Reading Research Unit and were generally held after school at a local centre. In 1961-62 when there were teething troubles in the administration of the experiment, the experimental group teachers had the opportunity of attending seven such meetings. In 1962-63, when the majority of schools in our research began the

---

*National Book League, (1965). *i.t.a. Books for the Teacher and the Child.* Available from the National Book League, 7, Albemarle Street, London, W.1.

first year of their experimentation, the initial problems had been cleared up. It was, therefore, not necessary to meet so frequently and only three meetings were held. Since then not more than two meetings each year have been provided.

In the design of this initial experiment it was recognized that while these meetings would be essential for the teachers using i.t.a., at the same time they might increase the Hawthorne Effect in this group although, of course, the importance of the Hawthorne Effect has yet to be demonstrated in educational research. Therefore, it was proposed that meetings of the control school teachers should also be held in an attempt to induce a similar increase in Hawthorne Effect (if any) in this group. It seems probable that these meetings of control group teachers have not been wholly matched in quality with those of the teachers using i.t.a. The purpose was obviously different. We tried, however, to show our concern for the teaching of reading in the control schools where t.o. was being used as usual. Each year a training college lecturer seconded to the Unit was given responsibility to meet this need to show special interest in the teachers in the control schools. Sometimes an outside speaker well-known as an expert in infants' school teaching methods was engaged to give an address followed by discussion. Sometimes the secondary training college lecturer conducted the meeting alone. For example, one series of control group meetings included an exhibition of the latest t.o. books and teaching apparatus for reading at the infants' school level, a lecture from the seconded training college lecturer and a discussion of children's books by the control group teachers.

These control group meetings were sometimes held in the afternoon, and in some cases lasted longer than the experimental group meetings. It seems inappropriate to attempt to compare these meetings of the two groups in terms of hours and minutes, but for the record we note that control meetings were held by the seconded training college lecturer on average twice each year, although these arrangements have not been uniformly successful. It seems fair, however, to suggest that although the experimental group leaders met rather more frequently, a reasonable demonstration of our interest in the work of the teachers in the control group was in fact provided.

After our main experiment had begun in 1961 we recognized that it would not be possible to control all of the many important factors in the experimental situation. We became concerned about two of the possible motivating influences noted by Southgate, i.e. visitors and publicity. Southgate comments that ' If equality of visitors and publicity was actually aimed at for experimental and control schools, it is doubtful whether this was achieved '. With this we agree, but it cannot be assumed that the presence of visitors is a tonic for teachers and children which automatically produces improvement in reading, because it could equally be a source of distraction and interference in the learning situation.

Large-scale experiments conducted in the normal social setting involve many complex problems of design, analysis and interpretation. These problems arise from the complexity of the educational situation in which there are many factors which are either known to affect attainment or which are believed to do so. Before giving a report on our second experiment we should first consider the incidental variables in research in i.t.a. By ' incidental variables ' we mean all those factors which may affect reading attainment, other than the one being studied. We may distinguish between those incidental variables which are a normal part of the educational situation and those which are associated with the experiment.

The first group includes, for instance, intelligence and home background of the pupils, teacher ability, methods of teaching and class size. The second group of incidental variables includes all those which arise from the experiment itself or from incidental circumstances surrounding the experiment and which may produce motivational effects not normally a part of the educational situation. Such 'extrinsic' motivational effects would not persist when the experiment and its surrounding circumstances cease to operate, and so are not a normal part of the educational situation.

The Hawthorne Effect, to which reference has already been made, falls within this second group. 'Hawthorne Effect' may be defined as an improvement in performance consequent upon participation in an experiment and independent of any other factors. It is presumed to be mediated by motivational changes. In our research in i.t.a., experimental and control groups both participate in the experiment, and insofar as both have been treated in the same way, there is no reason to suppose that there are any differential Hawthorne Effects.

However, it is possible that other extrinsic motivational effects have been produced in the i.t.a. experiment. The launching of the i.t.a. research required the simultaneous introduction of a new writing-system, and this innovation aroused considerable interest in the press, on television and in educational circles. These two factors, newness and public interest, *could* have had an effect on the performance of the experimental children, operating through changes in their own level of motivation or, more likely, through that of their teachers and/or families. There may also have been 'compensating' effects on the control group.

Except perhaps under laboratory conditions, extrinsic effects arising from newness and public interest are virtually outside the control of the educational researcher. It has been suggested that these factors could be controlled by adding a second experimental group using a second new medium or method, but it is doubtful if another new method would be equal to i.t.a. either in terms of perceived newness or the amount of public interest aroused. Probably the only way in which further information could be obtained on this point would be to study results obtained over time—not the usual kind of longitudinal study of pupils but a longitudinal study of groups of children entering the experiment in successive years, it being plausible to assume that the initial enthusiasm which it is claimed is aroused by newness and public interest is likely to diminish to vanishing point over the years.

In our original and main experiment begun in 1961 it is important to recognize that we have not had a free hand. We have always been dependent on teachers volunteering to join the experimental and control groups. At the beginning of the first experiment in 1961, the numbers of volunteers were small and as we were unable to use random sampling procedures, we attempted to establish experimental and control groups which were matched on certain school variables, which could be measured easily, i.e. size, urban/rural location, type of organization, pupil/teacher ratio, minimum age on entry, socio-economic level of the catchment area, and facilities of the school building.

We also attempted to match the two groups on certain pupil variables; age, sex, social class (as defined by father's occupation), Raven's Matrices score, Crichton Vocabulary score. All of the pupils were new entrants and they comprised the entire intake of a particular term or particular year depending on the size of the school. Experimental pupils began to learn to read with i.t.a., and control pupils with t.o. As we have indicated earlier in this paper, we tried to give equal attention to teachers and pupils in both groups, visiting schools of both types and holding research meetings for teachers in both groups, although our attempts were not entirely successful.

In brief, we were able to control or measure many of the variables which are a normal part of the educational situation. Others, including teacher variables and the less easily measured school variables, were not taken fully into account. Due to constraints beyond our control, we were unable to make any systematic assessment of teachers' attitudes or to deal very effectively with changes of attitude arising from certain extrinsic variables.

Aims and designs of the second experiment with i.t.a.

Our concern about the differences in the experiences of experimental and control groups and the lack of control over the teacher variable led us to launch a second i.t.a. experiment in September 1963. This second i.t.a. experiment controlled some of these incidental variables more rigorously than the earlier study had done. According to the design of this second i.t.a. experiment each participating school was to contain one experimental and one control running in parallel. On admittance, pupils were to be allocated randomly to these two classes, holding sex and age distributions constant. The two class teachers thus involved were to share their time between the classes in such a way that each teacher spent half of her time with each class. One of the teachers was to take the main responsibility for teaching reading to both classes, to one class with i.t.a., and to the other with t.o., using the same methods and giving the same amount of time to the teaching of reading with both classes. Both groups used the same series of basic readers. The experimental classes used *Janet and John* in i.t.a. while the control classes used *Janet and John* in t.o. The second teacher would, of course, have to do some reading work incidental to number and other activities but the reading teacher was to deal with activities directly specific towards the development of reading and writing skills. Both teachers were to keep the work done with one class parallel, in broad terms, with that done with the other class. This division of labour required that the teachers should change classes frequently, daily intervals being the largest that could be tolerated.

The advantage of this design is that most of the variables normally associated with reading attainment are controlled. School and pupil variables, in particular, are controlled effectively. Some important teacher variables, such as ability to teach reading and more general aspects of teaching ability dependent upon personality factors, are also controlled. Interactions between teacher and social group (class) are not controlled, nor are variables which are necessarily related to the writing-system. It may be that some teachers do best with i.t.a. while others get better results with t.o., or that a given teacher finds one class of children easier to teach than another. Neither of these possibilities is taken into account in the design. However, it seems unlikely that either of these effects would be as large as the general factor of ability to teach reading. In other words, a good teacher is likely to do well with either i.t.a. or t.o., though she may do relatively better with one than the other. It seems possible the teachers who have used i.t.a., after having been trained to do so, may show a general improvement in their performance as teachers of reading even when using t.o. However, the design of our second experiment ensures against a bias arising in this way also.

On the basis of this plan sixteen headteachers* agreed, presumably after consultation with their teachers, to take part in the experiment and to allow all of their 1963/64 entrants to participate. Over the whole school year this yielded a sample of 548 experimental and 554 control pupils, excluding those who left before the end of the first term.

In practice there have been some deviations from the original plan as far as the teacher variable is concerned. From the beginning we doubted that we should be able to find teachers who would adopt a procedure which flew in the face of current infant school practice. Once we had recruited sufficient volunteers it soon became clear that it would be difficult to maintain control of the teacher variable as rigorously as had been intended.

Preliminary discussions with headteachers confirmed our own opinion that it would be necessary to delay the 50/50 pattern of sharing teachers until the pupils had settled in school. The decision as to when the 50/50 pattern should be put into operation had to be left to headteachers. In some cases this happened within a few days of the beginning of term, but at the other extreme two schools took six weeks to implement the 50/50 plan completely.

In addition, other circumstances and events such as shortage of staff, absence of research teachers and festivities have led to the suspension of the 50/50 pattern for varying periods of time. Many of the teachers have expressed a dislike of moving from one class to another and it is probably this rather than any other reason which has led to lapses in some of the schools.

Figure 1 shows the extent to which the schools were able to maintain the 50/50 pattern of sharing teachers during the period with which this report is concerned (September 63-March 65). The information needed to calculate the actual ratios was obtained from the teachers who made retrospective estimates for the first year and who have completed daily timetables from September 1964 onwards. The ratios are averaged over the whole period.

Figure 1

Actual Ratio (September 63 - March 65)

	50/50	45/55	40/60	*poorer than* 40/60
No. of schools	2	9	2	3

---

*The Reading Research Unit wishes to express its appreciation of the help and co-operation of the L.E.A.s and headteachers involved in the experiment.

The three schools in which the reading teacher spent more than 60% of the time with one class and less than 40% with the other have been eliminated from the following analysis. The rest of this report refers only to the remaining thirteen schools.

Another point which was conceded in some schools was that for administrative purposes each teacher should be associated with only one class. In practice this has meant that a particular teacher always spends the first period of the day (usually about fifteen minutes) with one class and never with the other. During this period she takes registration and deals with other business. This procedure has been adopted by four of the thirteen schools. This has not been taken into account in the calculation of the ratios shown in Figure 1.

It seems possible that departures from the 50/50 pattern as it was envisaged in the original design may have favoured the experimental group children using i.t.a. In schools where an imbalance existed it was always the reading teacher who spent most time with the experimental class, and in those cases where one teacher became associated with one class for administrative purposes it was always the reading teacher who became so associated with the experimental class.

Turning now to other variables which the design was intended to control, no significant differences between the total experimental and control groups have been found on any of the following pupil variables: age, sex, social class (as defined by father's occupation), intelligence, vocabulary, measured as in the earlier project. Indeed, the distributions of control and experimental groups on all of these variables were very similar, as one would have expected.

Each teacher had been asked to use the same methods of teaching and to give the same amount of time to the teaching of reading in both her experimental i.t.a. class and in her control class. They were frequently reminded of the necessity of complying with this request. Five out of twelve teachers said that they did phonic work earlier with the i.t.a. class than with the t.o. class, but this was not seen as a change of method, because it was said to be response to the children's interest and development. As such, it was regarded as an essential part of the method used with both classes*.

Care was taken to accommodate parallel experimental and control classes in rooms with comparable facilities and to avoid similar class libraries. However, in three schools a shortage of i.t.a. books was experienced during the first half term. Although public libraries stocked i.t.a. books and parents were able to buy them, in general they were less readily available outside school than t.o. books were. Thus with regard to the supply of materials the t.o. classes appear to have had an advantage.

It had been suggested to us that the results might be contaminated if the control children were exposed to i.t.a. accidentally as might happen when experimental and control classes are situated in the same school. Consequently headteachers were asked to allow control children to enter experimental classrooms as little as possible. According to the teachers this has only happened very occasionally.

---

*This information was obtained from twelve of the reading teachers at the end of the first year. Among these were two from one large school where there were two pairs of parallel classes. No questionnaire responses were obtained from the remaining two reading teachers.

Steps were taken to reduce the differential operation of possible sources of motivation. When it was necessary to hold meetings for parents of experimental pupils to explain i.t.a. to them, parallel meetings were held for the parents of control pupils. Headteachers were requested to close the doors to visitors as far as possible and to make sure that those who could not be avoided were made to visit control as well as experimental classes. It is recognized that this does not guarantee that equal interest, approval, etc. was shown by visitors in both cases, any more than our meetings for parents ensure comparable attitudes on their part.

Outside agencies cannot be expected to be much concerned with the problems of educational research and their interest has generally been focussed on i.t.a. and the activities of the experimental classes. This is true of articles in the national and local press. We were, however, able to a large extent to prevent publicity for the schools actually participating in this second experiment. But incidents which we did not anticipate have been brought to our attention. For instance, at the end of the first year headteachers received without our knowledge or consent, a mailed questionnaire concerning their opinions about i.t.a., which was heavily biased towards answers favourable to i.t.a.

As far as extrinsic motivational effects operating through the teachers are concerned, this second experiment, as compared with the earlier main experiment of 1961, presents a rather different context in which such effects might be felt. Whereas in the first project the experimental and control teachers were usually working in different schools, generally not coming into contact with one another, in the second project the role of experimental teacher and that of control teacher are combined in one person.

While this arrangement effectively controls the factor of teacher ability there are other possible consequences which need to be considered. When the same teacher takes both the experimental i.t.a. class and the control t.o. class, she is able to make immediate comparisons of the progress of the two classes she is teaching. If i.t.a. does permit an easier beginning, a ‘ spiral ’ motivational effect may occur. Quicker progress in the i.t.a. class may produce positive attitudes in the teacher which enhance her work with i.t.a. Furthermore, she may become pessimistic about t.o., with a possible negative influence on her work in the control class.

Some evidence on this question is available from the questionnaires completed by the teachers at the end of the first year. Nine out of twelve reading teachers said, *in retrospect* that, at the beginning of the experiment, they were of the opinion that i.t.a. would be better than t.o. Two thought t.o. would be better and one said she had formed no opinion. ‘ Better ’ was taken to mean that the children would learn to read in t.o. more easily. Seven out of twelve teachers said that they had preferred teaching with i.t.a. One had preferred teaching with t.o. and four had no preference. Both of these items indicate that the majority of the teachers were biased in favour of the experimental i.t.a. group, by the end of the first year, if not earlier.

Ten out of the twelve teachers said that quite apart from their preference for the writing-system, they preferred teaching the experimental i.t.a. *class*. In other words, this preference is for a particular group of children rather than for a particular writing-system. Two teachers had no preference. The bias which appears to exist here is probably related to the special association which grew up between reading teachers and experimental i.t.a. classes which was described earlier. Nine of the teachers said that apart from their writing-system preferences, they found the experimental class easier to work with. Two had no preferences and one found the control class easier.

In an effort to counteract such expected preferences, teachers were reminded often of the need to give the two classes equal treatment and they may have succeeded in keeping their preferences covert. In considering this problem one should also take into account the possibility that any increase in motivation engendered by the i.t.a. teaching situation might spill over to the other in t.o. This seems to be true also of any improvement in teaching technique due, for example, to improved knowledge of the structure of English through experience with i.t.a.

The testing programme of this second experiment was the same as that used in the original main project. All of the testing was done by testers employed for that purpose. Testers were trained to give all tests, other than the Schonell Graded Word Reading Test for which they received mailed instructions only. The same tester tested both i.t.a. and t.o. groups within one school and the testing schedule was balanced to control time factors.

The chief object of the Reading Research Unit's experiments with i.t.a. is to determine whether t.o. is an important cause of difficulty for children in the initial stages of learning to read. For this purpose i.t.a. is the representative for simplified writing and spelling systems which may be contrasted with the complex and irregular t.o. writing system. To test the hypothesis that t.o. is a cause of difficulty in beginning reading it is necessary to compare i.t.a. attainments with t.o. attainments. Therefore the pupils in the experimental group are tested in reading the i.t.a. which they are learning, and their results are compared with those obtained from testing in t.o. the control group pupils who are learning to read with t.o. This applied to the first three tests given (Schonell Graded Word Reading Test, given twice, and the Neale Analysis of Reading Ability, Form C).

After this *both* groups are tested in *t.o.* Although the chief concern in all our research on i.t.a. is the comparison of i.t.a. attainments of children learning to read with i.t.a. with t.o. attainments of children learning with t.o., we have also to investigate the practical problem of transfer from i.t.a. to t.o. when children have begun with i.t.a. The Neale tests in Form A represent a measure of the effectiveness of transfer from i.t.a. to t.o. in the experimental group. However, this measure is clearly premature because one month before the test was given only 17% of the experimental i.t.a. group had been transferred by their teachers to t.o. reading materials, but *the whole* of the i.t.a. group was tested *in t.o.* This study had not yet progressed far enough to provide useful data on this second question.

Results of the second experiment

Results for the September entrants only are available as yet. January and April 1964 entrants have been tested and their results will be analysed later.

Table I summarizes the results of reading tests in which parallel versions of the tests were used, the ordinary t.o. versions being used with the control groups and an i.t.a.

version with the experimental groups. Table II shows results on the Neale Analysis of Reading Ability Form A when the ordinary t.o. version was used with both groups of subjects. These two tables give the mean scores of experimental and control groups within each school and the *t*-values obtained when these means are compared.

In many of the groups considered, the scores were markedly skewed due to the presence of small sub-groups of children who were far superior to the rest. (Such skewness is to be expected when tests are used at the bottom of the age range for which they are intended.) The mean, therefore, tends to be higher than the median value (i.e. *more* than half the children in each group lie below the mean value shown) and, in cases where the proportions of superior children are not balanced, the variances differ. To render the variances more comparable and to correct for skewness, it was, therefore, necessary in such cases to apply a logarithmic transformation to the raw scores. This procedure makes the matched samples statistically comparable, and enables the resulting *t*-values to be aggregated by a method outlined by Hald (1952).

The essential feature of Hald's procedure is that it takes account of the direction of differences. It therefore inquires whether the performance of the experimental children is, *on the whole*, superior to that of the control children and having regard to the fact that some of the comparisons favour i.t.a. whereas others point in the opposite direction. The resulting test statistic is referred to a *z*-table of percentage points of the normal distribution.

As Table I shows, the *z*-value exceeds the 1% level of significance in favour of i.t.a. on all of the tests appearing in this table:

1.	Schonell (first)	$p < .01$
2.	Schonell (second)	$p < .0001$
3.	Neale C Accuracy	$p < .0001$
4.	Neale C Rate	$p < .01$
5.	Neale C Comprehension	$p < .01$

We conclude that, *despite the conflicting evidence contributed by some comparisons that show the control children to be superior*, the balance of statistical evidence is significantly in favour of i.t.a. It need hardly be said that further investigations are required to elucidate the factors underlying the variation of differences between experimental and control groups and such investigation is currently under way. In this context, it should be noted that despite the similarity of the total experimental and control groups as far as pupil variables are concerned, differences on these variables may exist between the groups within schools.

More cases in which the performances of the control group is superior to that of the experimental group appear in Table II and the overall comparison is slightly, though not significantly, in favour of t.o. However, as indicated earlier, results on this test cannot be regarded as fair estimate of the performance of the experimental children after transfer since only 17% of them had been transferred to t.o. by their teachers at the time of testing.

Discussion of results from the second study

The chief trends which appear in the results of the second experiment are similar to those found in the original experiment. Pupils beginning with i.t.a. make more rapid

Table I

SUMMARY OF READING TEST RESULTS (i.t.a. versions of test used with experimental group)

TEST SCHOOL	Schonell 1st *Exp.* $\bar{x}$	Schonell 1st *Con.* x	Schonell 1st *t value*	Schonell 2nd *Exp.* $\bar{x}$	Schonell 2nd *Con.* $\bar{x}$	Schonell 2nd *t value*	Neale C Accuracy *Exp.* $\bar{x}$	Neale C Accuracy *Con.* $\bar{x}$	Neale C Accuracy *t value*	Neale C Rate *Exp.* $\bar{x}$	Neale C Rate *Con.* $\bar{x}$	Neale C Rate *t value*	Neale C Comprehension *Exp.* $\bar{x}$	Neale C Comprehension *Con.* $\bar{x}$	Neale C Comprehension *t value*
1	0.67	6.94	4.00**	4.82	14.62	1.81	7.2	14.23	0.71	7.9	15.62	0.55	1.00	3.69	2.04
2	9.44	9.42	0.01	26.00	21.37	0.84	21.16	24.05	0.52	33.53	41.53	1.16	7.32	9.63	0.41
3	5.31	4.00	0.80	25.69	12.58	1.06	20.62	14.50	1.17	21.54	19.33	0.61	7.46	5.58	0.94
4	16.00	6.19	2.40*	29.00	15.43	2.58*	20.56	12.19	1.65	36.88	28.21	0.93	7.63	5.50	1.07
5	9.89	2.33	1.69	23.67	8.57	1.56	18.00	8.57	2.05	12.78	12.00	0.19	4.33	2.86	1.11
6	6.37	1.69	1.15	17.27	6.21	2.59*	13.80	5.39	2.82**	12.86	6.97	2.34*	3.74	2.06	1.78
7	8.50	4.06	1.85	44.50	14.39	5.04***	34.42	15.06	3.23**	27.67	22.61	0.91	11.33	5.78	3.01**
8	8.65	4.70	1.42	23.17	11.24	2.99**	15.80	9.30	1.60	21.61	16.87	0.95	5.28	4.13	0.85
9	2.32	1.35	1.59	19.71	6.41	3.41**	15.19	7.53	2.52*	15.38	11.41	1.68	4.38	2.59	1.49
10	10.43	3.42	2.56*	18.68	8.81	3.02**	14.82	10.61	1.52	27.97	25.03	0.53	4.59	3.61	1.12
11	2.25	1.82	0.58	16.15	6.13	4.49***	14.73	7.23	1.99	17.67	8.94	3.29**	4.06	1.90	2.52*
12	4.06	5.11	0.36	16.94	11.60	0.86	11.65	10.50	0.24	12.53	18.00	0.08	3.24	4.20	0.69
13	14.33	8.29	1.30	23.14	15.22	1.48	18.71	16.28	0.36	40.29	26.89	0.48	6.57	4.83	0.82
z score		3.04			7.793			4.97			2.80			3.205	
sig. level		$p < .01$			$p < .0001$			$p < .0001$			$p < .01$			$p < .01$	

Table II

SUMMARY OF READING TEST RESULTS

(t.o. versions of test with both groups)

TEST SCHOOL	Neale A Accuracy *Exp.* $\bar{x}$	Neale A Accuracy *Con.* $\bar{x}$	Neale A Accuracy *t value*	Neale A Rate *Exp.* $\bar{x}$	Neale A Rate *Con.* $\bar{x}$	Neale A Rate *t value*	Neale A Comprehension *Exp.* $\bar{x}$	Neale A Comprehension *Con.* $\bar{x}$	Neale A Comprehension *t value*
1	3.30	19.00	3.50**	3.70	25.82	4.74***	1.00	4.00	2.73*
2	19.79	25.58	1.06	37.95	41.74	0.48	8.00	8.95	0.51
3	13.15	16.64	0.68	22.85	26.55	0.57	5.15	6.18	0.61
4	16.94	18.69	0.30	32.50	37.08	0.45	7.31	6.77	0.27
5	13.56	12.86	0.13	14.78	15.27	0.08	4.33	5.13	0.49
6	10.38	8.06	0.77	15.91	12.00	0.94	3.56	2.67	0.94
7	27.75	16.94	1.85	35.08	23.82	1.94	9.08	5.41	2.23*
8	9.94	9.06	0.23	21.06	25.75	0.65	3.67	3.69	0.02
9	8.83	8.06	0.28	12.78	12.28	0.16	3.00	2.39	0.63
10	12.66	12.36	0.10	27.06	27.69	0.12	4.97	4.33	0.64
11	9.82	8.93	0.33	16.30	17.70	0.35	4.24	3.80	0.41
12	7.47	12.20	1.05	10.35	15.20	0.87	3.20	4.50	0.83
13	14.86	18.89	0.56	29.43	30.67	0.13	6.29	7.33	0.42
z score		−.961			−1.496			−0.138	
sig. level		N.S.			N.S.			N.S.	

*—$p < .05$
**—$p < .01$
***—$p < .001$

progress and achieve superior word-recognition, speed and comprehension scores* in reading i.t.a. than do pupils learning and reading with t.o. However, the results obtained in this second study are by no means identical with previously analysed results from the original i.t.a. experiment which began in 1961. We have found that, in general, the differences between the experimental and control groups of the second experiment are smaller than those between the experimental and control groups of the original experiment. The scores of the experimental group of the second experiment tend to be lower than the scores of the experimental group of the original experiment. Preliminary analysis suggested that the same is true, though to a smaller extent, of the comparison between the two control groups. Space does not permit us to make a detailed comparison of the two sets of results in this article but this will be made in the Report to be published in 1966.

Meanwhile some questions may be posed as to why the superiority of the i.t.a. results in this second study are less marked than they were in the original experiment.

(*i*) Are there differences between the samples of the two experiments? Only the age distributions appear to vary but unless there is a very marked interaction between writing-system and age, so that older children benefit more from i.t.a., this factor alone could not account for the difference between the results of the first and second experiment.

(*ii*) Are there any differences between the two experiments in the teachers' attitudes towards the teaching of reading and its place in the infant school curriculum? There may be a tendency amongst the teachers in the second project to take a broader view of the educational process and to be in less of a hurry to get their pupils started on this one of the three R's. Insofar as the difference between the two experimental groups and that between the two control groups appears to decrease as the testing programme proceeds, the results support this hypothesis.

(*iii*) Are there any differences between the two experiments in the teachers' attitudes to i.t.a.? Were the teachers in the first experiment more anxious about the effects of i.t.a. and did they, therefore, take their pupils on faster? It is particularly notable that only 17% of the i.t.a. group had been transferred at the time of the first test in t.o. in the second experiment whereas the proportion was 40% in the first twenty schools to experiment with i.t.a. in the 1961 project. We have noticed an increasingly relaxed attitude towards the transfer stage as the use of i.t.a. has spread and teachers have become more familiar with the notion of transfer of reading from i.t.a. to t.o.

(*iv*) Have other incidental variables, such as teacher competence, visitor effects, local publicity and the less easily measurable pupil and school variables, which acted in favour of the experimental i.t.a. group in the original project been brought under more rigorous control in this second study as was intended?

(*v*) Is the 50/50 arrangement so unpopular and disturbing that reading achievement, in both i.t.a. and t.o. classes, is lower than it would be under normal conditions? Some of the teachers taking part in the second experiment believe that this is so. If this is true, although the comparison between the two media within the framework of the experiment is not invalidated, generalization of the results and comparison with those of the original experiment becomes problematic.

---

*In the Neale Analysis of Reading Ability the accuracy score sets an upper limit on the comprehension score. Therefore the i.t.a. children have more opportunity to achieve high comprehension scores.

Nevertheless, although there are these differences between the two experiments, and firmer conclusions must await a more detailed comparison of the results from them, this second i.t.a. experiment provides further, although still incomplete, evidence to support our previous tentative findings that t.o. represents a serious initial handicap to teachers and pupils engaged in the business of developing the basic skills of reading and writing in English-speaking countries.

References

Burt, C. (1962). Preface to Downing, J. *To be or Not Too Bee.* London: Cassell. (Title and text in Augmented Roman Alphabet.)

Downing, J. A. (1964). *The i.t.a. Reading Experiment.* London: Evans Bros. and (1965), Chicago: Scott, Foresman.

Downing, J. A. (1965). 'Current misconceptions about i.t.a.' *Elementary English*, May, 1965, pp. 492-501.

Gans, R. (1964). 'The initial teaching alphabet'. *Grade Teacher*, October, 1964, pp. 35 and 118.

Hald, A. (1952). *Statistical Theory With Engineering Application.* New York: J. Wiley.

Lewin, K. (1951). *Field Theory in Social Science.* New York: Harper.

Meade, E. (1965). 'The Initial Teaching Alphabet: A Beginning but Hardly an End.' Address given at the Second International Conference on the Initial Teaching Alphabet, Lehigh University, Bethlehem, Pennsylvania, 1965.

National Book League (1965). *i.t.a. Books for the Teacher and the Child.* London: National Book League.

Neale, M. (1963). *Neale Analysis of Reading Ability.* London: Macmillan.

O'Donnell, M. and Munro, R. *Janet and John.* Welwyn: Nisbet.

Putman, I. J. (1961). 'Learning to read: an experiment'. *Journal of Royal Society of Arts*, 109, 149-180.

Schonell, F. (1949). *Psychology and Teaching of Reading.* Edinburgh: Oliver & Boyd.

Southgate, V. (1965). 'Approaching i.t.a. results with caution'. *Educational Research*, vol. 7, no. 2.

University of London Institute of Education and the National Foundation for Educational Research in England and Wales, (1960). *Some Reasons Why we are initiating an investigation into the early stages of learning to read, when the matter to be read is printed in a special form alleged to be easy to learn and leading easily to a full reading skill.* London: University of London Institute of Education.

SCHOOL REFUSAL

by Margery G. Cooper

Institute of Education, University of Durham

COMPULSORY attendance at school for children, first made law in the Education Act of 1870, is a concept well established in this country. Responsibility for the maintenance of attendance is placed upon the parents, the child being prepared for and sustained in meeting the commitment. For the majority, adjustment to school is easy, bringing delight marred only by the odd difficulty or discomfort. For others occasional bids for freedom make possible tolerance of the authoritarian pattern—Tom Sawyer playing hookey personifies at least the secret wishes of countless small boys. Some children are unfortunate in having parents who are negligent and careless about the need for continuity of attendance. Truancy in its mild forms is viewed as an aberration in the total pattern of required attendance and Welfare Officers, through close attention to individual cases, keep the incidence as low as 1%. Only a small minority of cases remains persistent.

More recently, interest has focussed on a particular category of non-attender—children refusing to leave home or running away from school in order to seek refuge with the mother, eventually becoming involved in consistently long periods of absence. Irrational fear and anxiety make school an intolerable place and refusal to attend appears to have a phobic quality. Such children, suffering from ' school phobia ', are now seen to be in need of psychological help.

Thus, investigation into the full extent of the problem arising from unlawful non-attendance at school, known as School Refusal, is concerned with:

1. truants, either spasmodic or persistent, absenting themselves with or without the knowledge of their parents, for personal or social reasons;
2. school phobics, whose absence is the result of an emotional maladjustment known as school phobia.

The two broad categories, inevitably, have been found to overlap—a persistent truant may present many symptons of maladjustment—a truant may commit offences in order to sustain truancy and become delinquent—and there is great difficulty in diagnosis. Furthermore, liaison among the child helping agencies—the Children's Department, the School Welfare, School Psychological and Child Guidance Services, and the voluntary bodies—is over-involved. A child may pass through the hands of several officers rather than remain the concern of one, and treatment may lose its continuity. Investigation into the problem is hampered because of wide variations in diagnostic appraisal, closed avenues of treatment, and the lack of consistently detailed records.

Early investigations

The earliest investigations, notably that of Kline in 1897, attempted to explain truancy by linking it with a migratory impulse, ' a rebellion against suppressed activity and denial of free out-door life '. Such a simple diagnosis, however, could not account for the inexplicable urges of non-conformity which characterized the behaviour. In 1915 Healy produced evidence to link truancy with anti-social and unlawful conduct. His setting up, in 1909, of the 'Juvenile Psychopathic Institute' in Chicago led to the investigation of 1,000 juvenile court cases under a plan of research involving the techniques of academic psychology and social work. He emphasized the individual case approach using the services of the court officers and the personnel of educational centres. With an empirical approach he underlined the social causes of mental ill-health and deviant behaviour. Truancy, one form of deviant behaviour, was found to have as some of its determinants, poor parental control, broken homes and mental conflict. Evidence now existed to link the 'impulse to wander' with unlawful conduct. The importance of early detection and treatment by dealing with the basic and environmental causes, was stressed, lest truancy should become a precursor of crime.

Homes also were found to be supporting non-attendance at school. Many parents had not yet come to appreciate the advantages of freely offered education and kept the children at home to serve their own ends and for their own convenience. Girls who had reached a stage of helpfulness in the home stayed away from school in any domestic crisis, at times being almost permanently absent as in the case of a mother who needed support and companionship. Other cases of this kind were concealed under the guise of illness. Fretful children who preferred the warmth and comfort of home were supported by indulgent and over-anxious parents, and wilful children often beguiled them into accepting continued absence from school. It was not only Tom Sawyer who 'canvassed' his system.

Further work on delinquency and truancy

Burt, in *The Young Delinquent* (1925), questioned the 'inborn impulse to wander' by pointing out that all healthy children are curious and want to explore their surroundings. He showed in his work that the paramount conditions for truancy associated with delinquency were mental dullness, temperamental instabilities, defective family life in vicious homes and undesirable friendships with other delinquents. He found that 17.1% of boys and 4.1% of girls in his delinquent group had truanted. Truancy began to be more firmly associated with delinquency, and interest and investigation became focussed on anti-social conduct and rebellion against society in general. Truancy from the school, an agent of society, was only one of the multiple symptoms. The vast flow of work on delinquency—including, in America, that of Shaw and Mackay reported in *Delinquency Areas* and *Juvenile Delinquency and Urban Areas*, and the Gluecks' *Unravelling Juvenile Delinquency*; with, in this country, Birkett's *Delinquency in Aberdeen*—and the later work by Stott, *Delinquency and Human Nature*—revealed the delinquent as the social outcast. In 1932 Levy said in *On the Problem of Delinquency*, 'If a child who steals and truants is found to have learned all his lessons in life from a criminal father and an alcoholic mother in a poverty-stricken home in a delinquent neighbourhood, then one does not need to look far beyond the social causes'. The new concept of the delinquent as one who found himself in an emotional situation from which escape was impossible without aggressive reaction, was established.

The results of all the research on delinquency so brilliantly pioneered by Healy and Burt laid the foundation for future enquiry into the aetiology of personality problems and deviant behaviour in children. The work of Lowry (1936), Burlingham and Freud (1942), Levy (1943), Bowlby (1944), and Frielander (1947), on the influences of emotional relationships between parents and child and the emotional atmosphere of the home in determining the mental health of the child stressed the importance of the early relationships with the mother and the home group as the first socialising experience. The realisation of deep emotional factors, of anxiety created in the child's mind through guilt, jealousy, and repression of fears, encouraged the collating of evidence which verified and extended knowledge of the influences of home life. Defective family relationships, conflicts between parents over the upbringing of the child, unhappy marital relationships, inconsistency in family attitudes, lack of affection or outright rejection of the newly born, were all found to have an effect on the pattern of development.

Children thus made insecure in their relationships with home were not equipped to meet the consistent pattern of school, its disciplines and demands. New experiences of a substitute mother figure, of sharing equipment, of necessity to wait on another's pleasure, and of the turbulence of peer relationships could exacerbate feelings of non-conformity. The urge to seek some means of personal success, if expressed in a manner unacceptable to others, led either to complete rejection of school in running or staying away or to anti-social behaviour within the classroom, the inevitable punishment of which frequently strengthened feelings of resentment, unhappiness and anxiety.

Truancy was thus established as a discernible factor in larger syndromes of delinquent behaviour and mental ill-health. Unfortunately, research tended to concentrate on the more complex patterns. Investigators studying the isolated factor of truancy focussed their attention and interest on the part played by the school.

The school and the truant

In 1928 Dayton found that dislike of the teacher, feelings of inferiority, lack of recreational facilities after school hours, physical defects, lack of home discipline with a bad parental example and being kept at home by parents were more important factors in truancy than low intelligence. Educational retardation continued to be the main theme of investigation, and in 1931 McElwee, in a study of 110 New York river-front truants, found the medium amount of school retardation to be two terms. The children were truanting from their school as a means of escape from discouragement and feelings of inferiority caused by faculty grade placement.

In 1934 Brown stated that truancy was an attempt to preserve the child's ‘ sense of adequacy by escaping from a situation in which his inferiority undergoes constant magnification ’. Advocating psychotherapy for the children whose truancy was not caused by mental retardation or the influence of bad companions, he suggested that court action should be the last resort of the helping agencies.

In 1935 Kirkpatrick and Lodge reported a study of 752 cases of truancy referred to the Chicago Juvenile Court. The age-range of the children was from 6 to 17 years with a peak at 16 years: the girls tended to be older than the boys. The investigators found severe retardation in school; 7% were retarded 4 years, 16%—3 years, 25%—2 years, 22%—1 year, with only 7% at an average level for their age. The children were misfits in school and the study pleaded that the school curriculum should attempt to meet their individual needs.

Lippman, in 1937, urged that neurosis was a more potent factor in delinquency and truancy than had been previously recognized. In the same year O'Connor undertook a comparative study involving 119 truants. Intelligence and personality traits were investigated but no substantial conclusions were offered as to the relations between the two.

A further comparative study by Murphy in 1938, involving 50 truants and non-truants, showed the truants less intelligent as a group and more retarded in school achievement. They had, moreover, a larger number of physical handicaps. Murphy pleaded for special help in tutorial classes with emphasis on suitable manual rather than academic activities.

Further psychological and sociological studies

In 1939, Partridge, a Fellow in Child Psychology of Guy's Hospital, London, studied the isolated factor of truancy, fitting his cases into groups labelled, ill-disciplined, hysterical, desiderative or rebellious. One group of ten cases, well adjusted in school, were characterized by compulsion to leave school in order to see the mother. This group he labelled ' Psycho-neurotic ', victims of ' an emotional bond between parent and child the basis of which lies in undue attachment and over-protection '. Further studies with more of a sociological basis, notably those of Young (1947), highlighting emotional retardation and vicious home life, and Tyerman (1956), stressed what was by this time, reasonably well-documented evidence of unsatisfactory home conditions, inferior intelligence, social failure at school with consequent feelings of unhappiness, inferiority and frustration. Home conditions were seen as the major cause of the deviant behaviour.

The problem of ' school phobia '

Tyerman also referred to an associated problem of truancy—school phobia, which he described as ' a deep seated psycho-neurotic disorder, characterized by the terror of the child when at school '.

Research into the phobic aspect of School Refusal, pioneered by Broadwin in America '32, produced a flow of writing in journals of psychiatry and social work.

Description and Analysis. Broadwin studied only two cases of children absent from school with the full knowledge of the mothers. The children were suffering from deep seated obsessional neurosis and refusal to attend school was only one symptom of a multiplicity of personal difficulties. ' The child is absent from school for varying periods of several months to a year. The absence is consistent and at all times the parent knows where the child is. It is with the mother or near the home. The reason for the absence is incomprehensible to the parent and the school. The child may say that it is afraid to go to school, afraid of the teachers or say that it does not know why it will not go to school. When at home it is happy and apparently carefree. When dragged to school it is miserable, fearful, and at the first opportunity runs home despite the certainty of corporal punishment. The onset is generally sudden. The previous work and conduct had been fair '. During treatment the reasons emerging for the behaviour were fears of something happening to the mother. ' The reasons for the problem obtained in treatment from the child are meaningful in that they indicate the strong infantile love attachment to the mother, the intense sadistic or hostile attitudes towards her, which underlie the anxiety '. Running home brought reassurance and relief of anxiety. ' Absence from school relieved the anxiety and feelings of guilt and permitted the child to indulge himself further in infantile modes of existence by the undue contacts with the mother '. This brilliant diagnosis served as a guide for future investigators.

In 1941, Johnson, Falstein *et al.*, referred to this differentiation from truancy as ' school phobia ', highlighting the part school might play in the syndrome:

> ' Very early in this chain of events there enters a new factor, the school itself. When the teacher, as a more consistent disciplinarian, frustrates the child, she rouses his rage. Being less dependent on the teacher, who is a diluted form of the mother, the child's rage, inhibited toward the mother, can now find expression through displacement and the teacher in her milieu becomes the phobic object. To avoid the teacher and school is now the defence against being placed in the situation in which the overwhelming anxiety is aroused. '

In 1945 Klein, in America, expressed the opinion that most children experience some of the associated symptoms of school refusal in a mild form at one time or another, and stressed the necessity to recognize the school as a factor in the therapeutic programme.

Return to school ' at any level of school participation that the child can tolerate ' was stressed. Treatment, Klein asserted, is speeded up if the child maintains even fragmentary daily contact with the school. ' School anxieties may remain quite inaccessible as long as he is not in any contact with the object of his dread '.

Truants and Phobics. Warren, from the Institute of Psychiatry, Maudsley Hospital, studied from 1946-1948, differences between eight cases of ' special truancy ' and twelve ' typical truants '. The ' special ' cases were children suffering from a phobic reaction and were of higher intelligence, were more often only children, with subsequent spoiling, although with less trouble at school regarding progress and achievement in spite of personality characteristics of timidity, sensitivity and submission. Every child was over-dependent on the parents, with neurotic traits or past symptoms of emotional disturbance, and with the onset of the illness every child in the phobic group showed emotional disturbance in relation to the mother. Acute anxiety was prevalent. Defiant and disobedient, they did not run away from home or become involved in delinquent acts as did the truants.

Over a period of years research continued to highlight characteristics in the phobic behaviour.

Goldberg in 1952 made a useful distinction between the truants and the child who has a phobia. ' The phobic child is fleeing from and projecting his own aggression. Often he clings to the parents—at least they always know where he is and that he is in distress. The truant, on the other hand, evades his parent as well as the school authorities. Truancy arises predominantly out of social patterns and school phobia out of emotional ones '.

School Phobia and the Mother-child Relationship. In 1955 Bonnard, in a study of 7 cases, saw the problem of school phobia as one of a conflict of loyalties between school and its personnel and the home dominated by a psychiatrically disturbed mother. The child, loyally identifying himself with the home, absorbed the parent's abnormal attitudes and repressed the fears which arose as he met with the contrary standards of the outside world, especially in the school. Adjustment was only possible by a process of ' double think ', the strain of which resulted in breakdown when the situation unexpectedly demanded further adjustment as in change of teacher, class or school. In the same year Talbot commented on the increase of the problem. She investigated particularly the mother/child relationship, the neurotic nature of which revealed the inconsistent handling of the child by the mother, the ruling of the mother by the child and the feeling of guilt which predominated. The act of school refusal, involving an acute panic (agitated motive behaviour involving escape) brought shame and strengthened the mother's negative feelings. Talbot concluded, ' We would be interested in a follow-up study of school phobic children, in order to learn whether or not they are able to meet the next steps in development—marriage and employment '. Her work was given added impetus by Hitchcock who in the next year said, ' Many of the mothers who protest aloud their desire to have the child go to school, by their actions reveal a much stronger need to have the child stay at home '.

Separation Anxiety. In 1956 in a paper entitled *Separation Anxiety*, Estes *et al.*, writing from cases referred to the Mayo Foundation, Rochester, Minnesota, saw stubborn refusal to attend school as the presenting symptom of separation anxiety.

Therapy was conducted with the mother and child, bringing the conflicts of both into transference, partially satisfying dependency needs and allowing expression of hostility. Deliberate adjustments, so that the child returned to school, were made. Anxieties aroused permitted movement of the conflicts into the transference. Analysis of the conflicts was carried out and finally a situation was created which allowed both mother and child to resolve their dependency needs on a more mature level. Since the neurosis originated in the mother it was essential that she received treatment lest she involved another child in the ' same malignant relationship '.

In 1956, Coolidge *et al.*, from the Judge Baker Guidance Centre, Boston, Mass., attempted further to define the nature of the symptom. They said ' School phobia is a misnomer, it is a separation anxiety which occurs not only in early childhood, but also in later years '. Older children were found to be more deeply disturbed than younger.

School Phobia and Personality Disturbance. In a study of *Family Relations in the Development of School Phobia*, Waldfogel, Coolidge and Hahn took 53 cases from the Judge Baker Guidance Centre, Boston, of which 46 were in lower grades in the schools. In the cases of higher grade placings, however, there was evidence that there had been early symptoms. Among the younger children, the school phobia was more of a separation anxiety from the mother. They often remained motivated towards overcoming their separation anxiety and fear of school. In the older children the symptom was part of a wider character disturbance and the ' culmination of a long-term process '. In these cases the school phobia was an important part of a ' more diffuse picture of general fear of the outside world '. It is perhaps worthy of mention at this point that in their paper, ' Patterns of Anxiety. The Phobias,' Dixon, Monchaux and Sandler commented (1957) that it had long been established clinically that a patient while displaying one anxiety was, in fact, the subject of a much wider personality disturbance.

Later Investigations. In full agreement with the concept of school phobia as a clinical variant of separation anxiety, investigation in this country proceeded apace, the work of Model, of Nursten and of Kahn (1958) being followed by that of Morgan and of Hersov (1959), of Davidson (1961) and of Chazan (1962). Of these studies only that of Hersov had a statistical basis. He made a comparative study of children referred to Child Guidance clinics for treatment as a result of persistent non-attendance at school. Two groups of 50 children were studied, the first comprising ' those in whom this is a manifestation of a psychoneurotic syndrome which includes such conditions as anxiety reactions, obsessional disorders and hysterical reactions ', and the second, ' those in whom this behaviour is a manifestation of a conduct disorder—disturbed behaviour of a sort that invokes social or moral condemnation '. The main complaint of refusal to attend school in the first group and the feature of truancy in the second was investigated with regard to the environmental circumstances, the pattern of parent-child relationships and the personality and intellectual level of the child. A control group of children referred to child guidance for reasons other than the above was established. Results confirmed the existence of two groups with distinguishing features.

In the case of neurotic refusal to attend school the most common precipitating factor was a change of school or class. Early symptoms of anxiety could be noted in various forms during the first years of school life. Mothers of phobic children were commonly over-protective and indulgent, while fathers were passive and uninteresting. The family background had a high incidence of nervous disorder. The children themselves, while timid, fearful and inhibited in school had a high standard of work and behaviour. Maternal lack before and after the age of five was a significantly infrequent feature of the child's life, but there was a significantly high incidence of neurosis in the families. Children known as truants came from large, ill-disciplined families with little regard for social obligations. Changes in school had been frequent and the children, of inferior intelligence, produced poor school work. Hersov affirmed that ' Persistent refusal to attend school is not a clinical entity but one aspect of behaviour in a psychoneurotic syndrome '.

In the same year (1959) the 15th Inter-Clinic Conference of the National Association for Mental Health took as its subject, Truancy or School Phobia. In a paper, ' Truancy or School Phobia ', Green outlined the recent work done in this country on the two categories of non-attendance and made practical suggestions for treatment.

In 1962 the Association of Psychiatric Social Workers held a Symposium on School Refusal and in 1964 Kahn and Nursten published a complete review of the problem of School Refusal in their lucidly written book, *Unwillingly to School*. In essence, they plead for the consideration of problems of mental ill-health of which school phobia is seen only as one example.

Over the years, inquiry *within* the school has played only a minor part, particularly in the case of neurotic refusal. In 1961 Cooper selected two groups of phobics and truants, and attempted an assessment of the schools concerned, as well as teacher judgments of behaviour patterns of the children. These findings will shortly be published in this journal; similar studies might prove of considerable value.

The role of the school welfare officer

The role of the School Welfare Officer—historically of a primitive-enforcement character—has changed, and he is now more frequently recognized as essentially a social worker. The Officer maintains liaison between school and home, visiting school regularly to receive notice of non-attenders about whom school requires further information. He calls at the home, and if a medical certificate is produced, enters this information upon his attendance forms. Where reasons for absence are slight a kindly reference to the need for continuity of attendance is made, and on receiving reassurance from the parent he reports back to school. The spasmodic or incipient truant is detected at this stage when parents are shocked to learn that the child is not in school. Prompt action by parent and Officer can re-establish attendance.

The hard core of resistance comes from families in need of support either through improvidence leading to deprivation for the children, or through anti-social and unlawful conduct which includes refusal to recognize the necessity for the children to attend school. In the former case the Officer can supply clothing for the children, give advice to the parent and, at the same time, enlist help for the family from other supporting agencies. In the latter, where parents reject help and defy verbal warnings of the penalties for continued non-attendance (a night call is always made to present the situation to the father, the legal guardian), written warnings are given, along with invitations to meet a small committee set up for the purpose of offering advice based on an officer's report. If the invitation is declined and an improvement period has not re-established attendance, the full case is brought to the notice of the Education Committee who will provide authority to prosecute. Such resistant cases are eventually summoned to appear before a Magistrate's Court, and fined. The 1944 Education Act (Section 40) states that ' A parent found guilty either of failure to comply with an attendance order, or a failure to secure the regular attendance of a child is liable, for a first offence, to a fine not exceeding one pound, in the case of a second offence to a fine not exceeding five pounds, and in the case of a third or subsequent offence to a fine not exceeding ten pounds and/or to a term of imprisonment not exceeding one month '. It would appear that the law acts as a strong arm in support of the work of the Welfare Officer. In fact, the nominal fine of ten shillings, usually laid down by a Court not anxious to appear harsh and without understanding, comes as an anti-climax to the intensive work done on the case. The therapeutic value is low. Militating against the officer also are the delays between warnings and eventual prosecution which may be of six weeks duration, lulling the parents into a false sense of immunity from the law. Unfortunately there are also the inevitable postponements of hearings. The weekly Court of one hour is not necessarily exclusive to attendance problems and the cumulative effect of unheard cases brings further delay over which the Education Committee has no control.

If the child is truanting, defying home and all authority, the Officer acting in conjunction with the police, will trace the offender and personally escort him home. Persistent truants often have a known ' refuge '—the local market place, the cattle mart, or nearby large town. Others seeking solitude may prove more difficult to find, while unlawful acts of train jumping and sleeping out can become a pattern of behaviour for more serious cases. Diagnosis is often difficult but once the law has been broken the Welfare Officer is bound to route the case through the Juvenile Court where the Probation Service will take over.

If emotional disturbance is suspected before unlawful conduct has broken out and a Guidance Service is available, the Welfare Officer has some choice of action and may be instrumental in seeking for such a case psychological help in early stages. ‘ His choice of whether to route a case through the Courts or the Child Guidance Service can be decisive for the child's future ’ (*Unwillingly to School*, Kahn and Nursten).

In spite of Tyerman's (1956) warning, ‘ Truancy is the reaction of a personality to his environment and it requires, therefore, for its understanding, both a psychological investigation of the personality of the child who truants and a sociological study of his environment ’, psychological investigation appears not to be the first step in the treatment of truancy. It is only when persistent cases are referred to the Juvenile Court that the services of the psychologist are called upon. Moreover, the social therapy conducted in the early stages by Welfare personnel is in the hands of a body of men traditionally trained only in the hard school of experience, criteria for recruitment being concerned only with ‘ good ’ education and experience of family life.

The work of the School Welfare Officer still suffers from being regarded as ‘ legal enforcement ’ whereas it is now seen to be concerned with psychological/sociological approaches. Status and remuneration remain relatively low.

A course, newly instituted from a pilot scheme sponsored by the Educational Welfare Officers' National Association under the auspices of N.A.L.G.O. in 1961, now enables recruits and those in service to have opportunities for study and appreciation of the part social services must play in Child Welfare. Successful part-time study over three years leads to certification in Educational Welfare. Such a course could become compulsory and would do much to enrich the quality of treatment given to children. It should attract younger men to the service and make it possible to enrol women officers. Unfortunately it appears to be attracting only a very small percentage of men, though it seems likely that the qualification will be seen to be desirable for senior posts. Welfare Officers would welcome research into the sociological features of their work which would highlight successful remedial measures, suggest deterents and more useful forms of therapy.

The treatment of phobic and allied causes of truancy

Treatment for the child whose refusal to attend school has an anxiety basis is more varied as a result of the psychological approach. A large part may be played by the family doctor in referral, since early symptoms may have a psycho-somatic basis—the child complaining of feeling sick, of pains in the legs and stomach. If a doctor has not been consulted the Welfare Officer himself may recognize the symptoms, and in conjunction with the Headteacher, seek for the child attention from the School Medical Service. Cases with an anxiety basis may therefore be referred through the family doctor, the Headteacher or the School Medical Officer. School Psychological or Child Guidance Services, where these are available, give treatment at the hands of psychologist, psychiatrist, and psychiatric social worker. Such treatment is of course similar in scope and intention to that for any form of maladjustment and can be broadly divided into practical measures dealing with the environment, and psychotherapy, both of which are usually necessary. Adjustments within the school, change of school, hostel accommodation to relieve the child of the tension within the home, attendance at a day or boarding school catering for maladjusted children, or home tuition, freeing the child of the responsibility to attend school, are some of the more important practical measures. Psychotherapy, where ' the psychotherapist has to establish a relationship with the child in which the child can offer, and the therapist can accept and interpret both friendly and hostile feelings, and give in response to both, the same kind of gentle understanding ' (Kahn '58), seeks to lessen the underlying tension and anxiety the primary conflict. Psychologists have long stressed the need of treatment for both mother and child, allowing a period of time to elapse after breakdown during which the anxiety might be allayed. ' Care is needed in bringing the child to face his phobia, especially in the early panic phase ' (Morgan, 1958). That treatment should be prolonged and intensive, including in-patient hospitalization for more severe cases, is recognized (Hersov, 1959; Davidson, 1961) with residential schools and homes necessary when psychotic parents negate treatment. American studies have a different emphasis. The early assertion of Klein (1945), of the necessity to put the child after breakdown immediately ' in contact with the object of his dread ' is supported by later investigators condoning the judicious use of forceful measures and deliberate school adjustments admitting the possibility of having mother in school if necessary. The use of home teachers, conceded as a temporary measure in this country (Inter-Clinic Conference, 1959) is vigorously attacked by Eisenberg in America: ' Home teaching makes the situation too comfortable for the whole family and removes a major motivation for the change. By accepting the apparent inability of the child to attend as a real inability, it reinforces his regression. The insistence on attendance conveys to the child our confident expectation that he can accept and carry through a responsibility appropriate to his age '. The richer resources of the psychiatric reservoir in America are reflected in the use of second psychiatrist working with the mother, the emphasis on whole-family treatment (Waldfogel, 1956), and the employment of a psychiatric team within the school for research purposes. In contrast, it remains a fact that some Local Education Authorities in England and Wales, although empowered by the 1944 Act to establish Child Guidance and School Psychological Services have not yet done so. Many clinics, moreover, are functioning under conditions of severe strain because of the shortage of trained personnel. All clinicians, particularly in the provinces, are aware that at the moment they are only able to touch on the fringe of problems besetting children of school age. Finally, exclusion from school must be mentioned as a possible conclusion to cases which do not respond to therapy, however intensive. A higher incidence of these is found among adolescents, the relationship between age and severity being agreed upon by all investigators. Prognosis with older children is unfavourable, the possibility of schizophrenia being admitted.

The function of the school and the teacher

Research into the part played by the school in the problem of School Refusal has been slight. Earlier studies had stressed the need for adjustments in the curriculum to meet individual levels of retardation among truants but the establishment of remedial measures in the total problem of backwardness is now a recognized feature of school life. Matters of organization and methods of teaching in the schools of phobics and truants were investigated by Cooper (1961), but the 97 schools involved in the survey exhibited a full range of practice. Difficulties in controlling the variables inherent in human situations prevented this part of the study from having a comparative basis, but there was some evidence that schools describing themselves as ' informal ', i.e. where full acceptance of the theory of individual differences and rates of learning resulted in a permissive climate and where the curriculum provided creative opportunities for child involvement and learning, had fewer cases. That all schools were concerned about such children was shown in the considerable part they played in referral and in the adaptive measures many Headteachers instituted.

Within the classroom itself, there are danger signs which the observant teacher may detect. Among the younger children, disruptive behaviour, wandering from one task to another, ceaseless touching and pulling about of apparatus, violent and out-of-proportion temper tantrums, refusing to move and standing rooted to one spot, may all be indications that the child is not finding satisfaction in his school life, and is looking for an opportunity to reject it through truanting. A quieter, less overt type of behaviour, such as the lack of participation in classroom life, a hanging back and unwillingness to give of self, an over-dependency on the teacher and a seeking for physical proximity may be among the first signs of anxiety.

Among older children recurrent sore throats, headaches and mystery pains requiring short but frequent absences should arouse the teacher's suspicion. It is possible that many acts of spasmodic truancy are never fully investigated when the child returns to school with a note from an indulgent parent offering a weak excuse for absence. General lack of concentration, boredom, acts of defiance, petty thieving, loss of all spontaneity, or irritability and explosions of temper and attempts at physical violence should be noted. Many children with a record of spasmodic truancy are reported to be ' bad-tempered ' with their peers.

The feature of the sudden irruption of breakdown so often reported in the case of school phobia lends emphasis to the psychologist's plea that perhaps the greatest danger to personal development is withdrawal from human contact, the first sign of which is extreme shyness. There is evidence (Waldfogel, 1957) that phobic children tend to rely on magical thinking processes more than do their peers and retreat from normal social situations. Such constricted personality development in which emotions and phantasy life are affected by anxiety about the mother at home is a well-documented characteristic of the phobic child.

For her part the teacher can attempt classroom therapy by seeking to establish a good relationship with the child which will be likely to support him if he is experiencing difficulties with his peers. It should be possible for every child to experience satisfaction and success through the order of classroom affairs since the teacher has power to manipulate experiences likely to please and gratify. The extra thought and care, the bright smile and appropriate word of encouragement can become the concern of Staff and Headteacher once the problem is reported upon.

Conclusions

If the truant is to be expected to accept school and the phobic to be put ' in contact with the object of his dread ' more research must be done to establish which parts of school life are most difficult for them to bear.

Well documented evidence of the times of irruption, beginning school, change of school or class, suggests that research is needed into means of alleviating stress at these critical times, as well as into the structure of the psycho-social climate of reception classes at primary and secondary stages. A more positive attitude could be adopted by administrators, Headteachers and teachers to help minimize the effects of change and transfer. The benefits of staggered entry of small groups of 5-year-olds, of half-day attendance in the first year, should be judged in the light of evidence from countries with more permissive attitudes towards the attendance of school beginners. Research could help establish minimum and maximum periods needed for adjustment of differing behaviour patterns. The relief of pressure and tension from junior school life through the elimination of competitive experiences, the association of infant and junior teachers with common principles, aims and methods, the extension of primary education to a later age and more intercommunication among schools, could do much to lessen tension and anxiety among children.

Liaison between school and home should be maintained by trained social workers with their base in school. Research into the overlapping of helping agencies may point to the necessity of a school-based Welfare Officer. Work study methods could throw light into record-keeping patterns to facilitate and standardize the work. Classroom therapy based on an understanding of children's problems of adjustment would appear to make demands on the personal stability of the teacher. More research is needed in the field of teacher recruitment in order to discover the best ways of revealing personality strengths among those seeking the opportunity to teach. The whole pattern of education for teaching must acquire emphases which will enable teachers to have insight into children's problems. Knowledge and ability to shape healthy attitudes could enable teachers to be accepted as members of a team responsible for the mental health of children. Research also is needed into a wider interpretation of techniques of teaching to include socialising role-play, where teachers can be helped to establish for their children co-reactive and security tendencies. The special needs, common to both truants and phobics, appear to be for help in strengthening the self-image.

Underlying all is the need for smaller classes, with more ancillary and auxiliary help. Schools should have ready access to well staffed Psychological and Guidance Services so that teachers may have advice and children receive specialist help in the earliest stages of adjustment problems.

Finally, little is known of the later lives of children who have shown reluctance to accept responsibility of attending school. It seems possible that a pattern of dis-function could be established over the school years to irrupt at intervals and disorganize adult life. More research is needed to help the development of a national positive attitude to mental health in children. It would seem that this could best be achieved by focussing therapy on young children in a situation in which they all must participate, viz. school. If it is at all established that behaviour disorders have their origin in the early years of life, school, seen as the complement to home, must be where all children can feel happy and safe.

Bibliography

Birkett, N. T. (1950). ‘Aspects of Juvenile Delinquency in the City of Aberdeen’. Ed.B. Thesis, University of
Aberdeen. *Brit. J. Educ. Psychol.*, 22.
Bonnard, A. (1952). ‘School Phobia. Is it a Syndrome?’ *Prem. Con. Mon. de Psych.*, vol. VII.
Bonnard, A. (1959). ‘School Phobia’. Paper presented at a meeting of Ass. of Child Psychology and Psychiatry.
Bowlby, J. (1951). *Maternal Care and Mental Health.* Geneva: W.H.O.
Bowlby, J. (1956). ‘The effect of mother-child separation—a follow-up study’. *Brit. J. Med. Psych.*, 29.
Broadwin, I. T. (1932). ‘A Contribution to the Study of Truancy’, *Am. J. Ortho.*, 11, 3.
Brown, F. (1934). ‘The Truant Child’. *School and Society*, 40.
Brown, K. M. (1963). Symposium on School Refusal. Association of Psychiatric Social Workers, Feb. 1963.
Burt, C. L. (1925). *The Young Delinquent.* London: U.L.P.
Chazan, M. (1962). ‘School Phobia’, *Brit. J. Ed. Psych.*, 32, 209.
Coolidge, J. C. and Hahn, P. B. (1956). ‘School Phobia: Neurotic or Way of Life’, *Am. J. Ortho.*, XXVII.
Coolidge, J. C., *et al.* (1960). ‘School Phobia in Adolescence. A Manifestation of Severe Character Disturbance’,
Am. J. Ortho., XXX.
Davidson, S. (1961). ‘School Phobia as a Manifestation of Family Disturbance’, *J. of Ch. Psychology and Psychiatry.*,
vol. 1, no. 4.
Dayton, N. A. (1928). ‘Mental Deficiency and Other Factors that Influence School Attendance’, *Mental Hygiene*, 12.
Dixon, J. J., Monchaux, C. and Sandler, C. (1957). ‘Patterns of Anxiety. The Phobias’. *Brit. Journal Med.*
Psychology, XXX.
Eisenberg, L. (1959). ‘School Phobia. A Study in the Communication of Anxiety’, *Am. J. Psychiatry*, 114, no. 8.
Estes, H. R., Haylett, C. H. and Johnson, A. L. (1956). ‘Separation Anxiety’, *Am. J. Psychotherapy*, X, no. 4.
Fornwald, R. J. (1947). ‘Towards an Understanding of Truancy’, *School Review*, Feb.
Freud, S. (1926). *Inhibitions, Symptoms and Anxiety.* London: Hogarth (1948).
Freud, S. (1933). *New Introductory Lecture.* London: Hogarth (1949).
Glueck, S. and Glueck, E. (1950). *Unravelling Juvenile Delinquency.* New York: New York Commonwealth Fund.
Goldberg, T. B. (1952). *Factors in the Development of School Phobia.* Smith Coll. Studies in Soc. Work, 23.
Haase, E. (1932). Zeitschr f.pad Psychol., XXXII.
Hall, M. B. (1951). *Psychiatric Examination of the School Child.* London: Arnold.
Healey, W. (1915). *The Individual Delinquent.* London: Heinemann.
Hersov, L. (1959). ‘A Psychiatric Study of School Refusal’. M.D. Thesis, Inst. of Psychiatry, London.
Hitchcock, A. B. (1956). *Symbolic and Actual Flight from School.* Smith Coll. Studies in Social Work 28.
Inter-Clinic Conf. 15th (1959). *Truancy or School Phobia.* London: National Ass. of Mental Health.
Jacobsen, V. (1948). *Influential Factors in the Outcome of Treatment of School Phobia.* Smith Coll. Studies in Soc. Work 18.
Johnson, A. M., Falstein, E. I., Szurek, S. and Svendsen, M. (1941). ‘School Phobia’, *Am. J. Ortho.*, XI.
Kahn, J. H. (1958). ‘School Refusal’, *Med. Officer*, vol. 100.
Kahn, J. H. and Nursten, J. (1964). *Unwillingly to School.* London: Pergamon Press.
Kirkpatrick, M. and Lodge, T. (1935). ‘Some Factors in Truancy’, *Mental Hygiene*, no. 19.
Klein, E. (1945). ‘The Reluctance to go to School’, *in*: *Psychoanal. Study of the Child*, 1.
Kline, L. W. (1897). ‘Truancy as related to the Migratory Instinct’, *Ped. Sem.*, 5.
Lippman, H. (1937). *American Journal Ortho.*, VII.
Maberley, A. (1945). ‘Personality of the Problem Child’, *Brit. J. Ed. Psych.*, 16.
Markey, O. B. (1941). Discussion of ‘School Phobia’ by Johnson, *et al.* *Am. J. Ortho.*, XI.
McElwee, E. W. (1931). ‘A Study of Truants and Retardation’, *J. Juv. Res.*, 15.
Mercer, M. (1930). *Journal Juvenile Research*, XIV.
Model, A. N. and Shepherd, E. (1958). ‘The Child who Refuses to go to School’, *Medical Officer*, vol. 100.
Morgan, G. A. V. (1959). ‘Children who Refuse to go to School’, *Medical Officer*, vol. 102, no. 18.
Murphy, W. C. (1938). ‘A Comparative Study of 50 Truants and Non-Truants’, *J. Juv. Res.*, 22.
Nursten, J. P. (1958). ‘The Background of Children with School Phobia’, *Medical Officer*, vol. 100.
O’Connor, Z. (1937). *Psychological Bulletin*, XXXIV.
Partridge, J. M. (1939). ‘Truancy’, *J. Ment. Science*, vol. 85, Jan. 1939.
Shaw, C. R. (1929). *Delinquency Areas.* Chicago: University of Chicago Press.
Shaw, C. R. and Mackay, D. H. (1932). *Juvenile Delinquents and Urban Areas.* Chicago: University of Chicago Press.
Stott, D. H. (1950). *Delinquency and Human Nature.* Dunfermline: Brit. Carnegie Trust.
Stott, D. H. (1952). *Saving Children from Delinquency.* London: U.L.P.
Suttonfield, V. (1954). ‘School Phobia: A Study of Five Cases’, *Am. J. Ortho.*, XXIV.
Talbot, M. (1955). ‘Panic in School Phobia’, *Am. J. Ortho.*, XXVII, no. 2.
Tyerman, M. (1955). ‘A Research into Truancy’. Ph.D. Thesis, University of London.
Van Houten, J. (1948). *Mother-Child Relationships in Twelve Cases of School Phobia.* Smith Coll. Studies in Soc. Work.
Waldfogel, L., Coolidge, J. C. and Hahn, P. B. (1956). ‘The Development, Meaning and Management of School
Phobia’, *Am. J. Ortho.*, XXVII, no. 4.
Warren, W. (1948). ‘Acute Neurotic Breakdown in Children with Refusal to go to School’, *Archive of Diseases in*
Children, vol. 23.
Williams, H. D. (1927). *Journal Applied Psychology*, XI.
Young, A. J. (1947). ‘Truancy. A Study of Mental, Scholastic and Social Conditions in the Problem of Non-
attendance at School’, *Brit. J. Ed. Psych.*, XVII.

'METHOD-REVERSION': THE PROBLEM OF SUSTAINING CHANGES IN TEACHER-BEHAVIOUR

by J. D. Williams

National Foundation for Educational Research

IF it is important for the research worker to discover ways of improving education, it is also important that he should investigate ways of implementing improvements. Of the many obstacles to implementation, the most obvious and frequently discussed is that which is constituted by failure to communicate research findings to the teacher. It is perhaps not difficult to see that beyond the communication barrier lies a further obstacle constituted by the teacher's understandable reluctance to exchange the familiar for the untried. But, from where we at present stand—having made little progress in surmounting either of these obstacles—it is rather more difficult to appreciate a somewhat less-immediate member of this formidable series: the problem of *sustaining* changes, once these have been made.

In the present article we shall describe some aspects of the course of events that followed upon an attempt to introduce two new methods of teaching primary school mathematics—the Cuisenaire and the Dienes M.A.B.—into first- and second-year junior school classes. Most of our observations seem to be consistent with the assumption that while the teacher may agree—even enthusiastically—to adopt a new teaching-method, and may go to considerable lengths to acquire the skills necessary for operating this method, he nevertheless may tend to return to patterns of teaching-behaviour that are more appropriate to the method that has been superseded. This tendency could be called *method-reversion*. We shall attempt to illustrate some of the forms that method-reversion may take, to indicate some ways in which the kind and degree of its manifestation may be determined by characteristics of teaching-methods, and to suggest some approaches to its further investigation.

Our observations are a by-product of an experiment,* the main purpose of which was to compare the Cuisenaire and Dienes methods with others. In this experiment these methods were introduced into schools in which they had not previously been used, and teachers were given complete freedom to decide what kinds of use should be made of them.

The new methods

In some respects it is artificial to separate the aims and policies of a teaching-method from its techniques, but for purposes of simplicity we shall do so in the case of the Cuisenaire and the Dienes.

The Cuisenaire Technique. Wooden rods, from one to ten units in size, are used for demonstrating mathematical and numerical relations. These rods are not segmented, and therefore the pupil is not tempted to ascertain their size by counting; neither, in the initial stages of learning, are they referred to by the usual numerals—another feature that dissuades the pupil from employing counting procedures. Reference to the rods is made in terms of their colours; by this means mathematical relations can be described without specification of absolute numerical values. By combining, separating and comparing rods, pupils parallel many mathematical operations, and subsequently learn to carry these out with numbers, using the more conventional arithmetical notation.

---

*See Williams, J. D., 'Some problems involved in the experimental comparison of teaching methods'. Educational Research, Vol. VIII No. 1.

The Dienes M.A.B. Technique. The apparatus of the Dienes Multibase Arithmetic Blocks models five different number-systems. Dienes holds that if the pupil is acquainted with a variety of number-systems he will more easily be able to abstract the essential features of the concepts involved in number-systems. For each system there are pieces of wood of four different sizes: unit cubes; 'longs' (a number of units on end); 'flats' (a number of longs side-by-side); and 'blocks' (large cubes composed of a number of flats). In the case of the decimal number-system, these sizes would be 1 unit, 10 units (tens), 10^2 units (hundreds) and 10^3 units (thousands). In that of the base-3 system, they would be 1 unit, 3 units, 3^2 units and 3^3 units.

Using this apparatus, pupils perform tasks that not only show them how to compute, but also enable them to abstract the concepts underlying computation in general, and particularly those underlying the conventions of positional notation.

Policies. The proponents of both methods give importance to certain aspects of learning: the understanding of mathematical principles; the performance of concrete operations; play and discovery; opportunity for individual styles and rates of learning. However, while Cuisenaire discourages counting, Dienes does not, and while Dienes stresses the importance of variety of experience, Cuisenaire does not.

Some kinds of reversion

Teachers tended to revert in two different ways to approaches more suitable to earlier methods: either they restricted the use they made of the new method, or they used it in a modified form. Generally speaking, restriction of use occurred more often in the case of the Dienes, while the Cuisenaire method was more often modified.

Restricted Use. It was possible to isolate four species of this kind of reversion:

(i) Use of the method only at certain ability-levels: often, only bright pupils used the Dienes; on the other hand, bright pupils in Cuisenaire classes sometimes spent very little time using the apparatus.

(ii) Withdrawal of the method from use in the teaching of pupils who are poor readers; these pupils, particularly in the first year of the junior school, found it difficult to read the Dienes task-cards and in some cases they were taken off the Dienes method altogether.

(iii) Confinement of the method to special periods. Perhaps the new method would be used only once a week, while for the other mathematics periods more conventional methods would be used. The Dienes method was often used in this way.

(iv) Use of the method for special purposes only. Sometimes the new method was used only to illustrate difficult or new processes.

Modification. The forms that modification took were too numerous for detailed description, but they can be placed into two categories:

(i) Modifications of technique. The Dienes method seemed to be the more vulnerable to these. Most frequently this kind of modification took the form of teaching only in the decimal system.

(ii) Modifications of policy. The Cuisenaire method was the greater sufferer from these. It was not at all unusual for Cuisenaire teachers to lapse into class teaching and a use of conventional numerals which would be regarded by the devisers of the method as premature.

Some general causal factors

Before describing those particular aspects of the Cuisenaire and Dienes methods which produced reversion, we shall list some general factors, which might be expected to operate in the case of any new method.

(i) Shortage of materials. Where there was a shortage of the somewhat expensive apparatus involved in the operation of the new methods, there was (not surprisingly) a tendency to use the method in a partial, supplementary capacity (for purposes of illustration, enrichment, etc.). We did not establish whether shortage of apparatus was the cause of this kind of restriction, or its effect.

(ii) Application of criteria that were more appropriate to previous methods. It is tempting to assess the effectiveness of a teaching method in terms of the pupil's acquisition of specific skills rather than in terms of his grasp of general concepts, for, in many respects, the former of these accomplishments is easier to ascertain. Where the teacher has been accustomed to assessment in these tangible terms, the temptation must be even greater. There were indications that some of the teachers who were dissatisfied with the progress their pupils were making under the Dienes method, and who, for this reason, restricted their use of this method, had fallen to this temptation.

(iii) Time-budgeting. This is a factor that is likely to cause disappointment with any newly-introduced method. Until the teacher has covered a course at least once with the new method, he cannot hope to know how to distribute emphasis over various topics. Where he errs on the side of underestimation of the time that should be allowed for topics that come later in a course, he is likely both to be disappointed with the new method and to resort to more familiar techniques of teaching in order to make up for lost time.

(iv) Reduction of stress. Many aspects of the new methods conspired to place the teacher under a certain amount of stress; teachers often reacted to this by modifying or limiting the method. The use of apparatus in itself causes a certain amount of organizational difficulty: it must be taken out and put away; it clatters; it gets lost; in the case of the Dienes, there must be continual switching among children from one set of apparatus or cards to another. The policy of encouraging play and discovery is easier to recommend than to effect: conscientious pupils continually encounter obstacles to further progress, over which they need to consult the teacher, while weak pupils often use their freedom for making noise rather than progress. The policy of individual learning entails that different pupils are at any one time unlikely to be tackling the same part of the course, and that the teacher, apart from having to deal with a great range of learning problems in any one lesson, must abandon the economical device of class-teaching, and adopt in its place the very strenuous procedure of explaining each new step as it arises for each child or group of children.

(v) Persistence of behaviour that had developed as a response to early stress. There are indications that during the particularly stressful initial stages of using the new method, the teacher had sometimes been forced to return to ways of teaching that were more familiar than those required by the new method, and that he continued to use these even after he had managed to master this method.

Method-determinants of reversion

In terms of kinds of reversion our two new methods differed markedly. As pointed out above, modification tended to occur in the case of the Cuisenaire, while restriction tended to occur in that of the Dienes. The most conspicuous differences between these methods were in terms of return to previous teaching-policies. This kind of reversion was much more apparent in the case of the Cuisenaire. This particular difference between patterns of reversion might well be explained in terms of differences between the two methods.

Reversion to Symbolism. An early reduction in the use of apparatus, and a corresponding increase in work with symbols, tended to occur in the case of the Cuisenaire method, but not in that of the Dienes. Three factors may have contributed to this difference:

(i) The content of the mathematics course used with the Cuisenaire method more closely parallels that of the courses used with conventional teaching methods. In teaching the same topics as he had previously taught, the teacher may well have been tempted to use the same teaching techniques as he had previously used. It is unlikely, however, that the teacher would have had earlier experience in teaching the topics involved in the Dienes course, and correspondingly unlikely that he would have been able to find, in his conventional repertoire, techniques that were appropriate to this subject-matter. For example, before the Dienes teacher could employ 'table-chanting' he would have to modify the chants to cope with a situation in which the number-'facts' would be known to vary in relation to different number-systems.

(ii) While the colour-notation used in the Cuisenaire method differs from the conventional notation, it can almost as easily be divorced from reference to concrete situations, for colour is only arbitrarily related to the aspect of concrete situations which determines the nature of their mathematical 'behaviour'—that is, size. The Dienes method is to some extent protected from this schism between symbols and apparatus, for, in this case, all symbolic reference must include a specification of the base of the number-system involved and must be in terms of the shapes of the apparatus ('longs', 'flats', etc., instead of 'tens', 'hundreds', etc.). Partly because it is easier to think of the concrete situations to which this clumsy notational system refers than to learn a vastly-increased number of symbolic relations, but partly also because the apparatus shapes are easily-visualized and imply size relationships, it is probably more difficult for Dienes pupils to lose touch with the concrete.

(iii) Whereas the textbooks of the Cuisenaire method tend generally to present the pupils with tasks in terms of number-symbols and in a form that seldom involves reference to concrete situations, the Dienes cards tend to avoid excessive use of number-symbols, and to present tasks in terms of manipulations of apparatus.

Reversion to Teacher-centred Policies. Cuisenaire teachers tended to return to class-teaching of a kind that was incompatible with the full pursuit of the policies of fostering learning by discovery and allowing individual rates of progress. This species of reversion was almost entirely absent in the case of Dienes teachers. It is easy to find explanations for this difference:

(i) It is much cheaper and more convenient, in the case of the Cuisenaire apparatus, to provide each child in a class with a complete set of the components required for any operation he might be expected to perform. This facility means that, at any one time, all pupils can be performing the same operation with the same components—a synchronization which would not be practicable in the case of the Dienes apparatus. Again, whereas the Cuisenaire tasks are provided by means of one textbook for each child, the Dienes tasks are provided by means of a set of cards that can be adequately shared among the children of a class only if different children perform different tasks. This is another factor that facilitates synchronization of activity in the case of the Cuisenaire method but inhibits such synchronization in that of the Dienes. Where the activities of the children of a class can be synchronized, the teacher may be tempted not to make allowance for the different rates and kinds of learning of different individuals.

(ii) The Cuisenaire colour-notation constitutes another facility which might tempt the teacher to revert. This terse notation, unlike the clumsy notation of the Dienes, enables the teacher to communicate with the pupil at a distance. Where this is possible, the teacher is able to prescribe and monitor each pupil's manipulation of the rods orally, and, what is more, to do this for the whole class at once.

Instructions and Theory. In addition to the above differences, there are two further differences between the Cuisenaire and the Dienes methods which might be expected to prevent reversion to conventional teaching policies in the case of the latter, but not in that of the former:

(i) The Dienes cards go much further towards ensuring that pupils learn by discovery than do the Cuisenaire textbooks.

(ii) Dienes has provided the teachers with a clear and convincing theoretical framework which enables them both to rationalize his teaching policies and to derive applications of them to new situations as these arise.

The comments in this section may suggest that the Cuisenaire method is much more liable to reversion than the Dienes. But it must be remembered that the Dienes method often resisted modification only at the expense of being severely restricted or completely abandoned.

Conclusions

In the foregoing an attempt has been made to illustrate, by means of two methods of teaching primary school mathematics, some of the features and causes of the teacher's tendency to revert from patterns of behaviour required by a new teaching-method to patterns that are more appropriate to a method with which he is more familiar. Method-reversion cannot be regarded as a trivial phenomenon, for it clearly poses serious problems both for those investigators who research into teaching-methods and for those teachers who wish to implement the findings of such research.

In the present case this phenomenon has manifested itself as a side issue of an experiment whose main preoccupations are in other areas; however, our observations may have helped to indicate some lines along which we might make a direct approach to investigating the phenomenon and to developing strategies by means of which it may be controlled. By way of making a start, perhaps we could tackle some of the following questions:

What kinds of stress accompany the introduction of a new method, and how might these precipitate reversion? Which features of various techniques, policies and subject-matters produce stress and precipitate reversion? How do we remove revertive behaviour which, although precipitated in the early, more stressful, stages of use of the method, has persisted into later stages? What can be done in the early stages to improve time-budgeting? Why, in assessing the effectiveness of a new method, do some teachers employ outmoded criteria? Where methods (like the Dienes) resist modification, must they also invite restriction or abandonment? What roles can a theoretical framework play in sustaining changed teacher-behaviour? How can safeguards against reversion be built into the presentation of new methods? How is reversion influenced by teacher-factors (such as age, experience, etc.) and aspects of the teacher's environment (such as social support, etc.)? What kinds of tactic could the teacher employ as a means of sustaining changes in his approach to teaching?

At a time at which attempts are being made to effect dramatic changes in teaching methods and curricula, we are provided with both the need and the opportunity to answer such questions as these.

Research Note:

THE MINNESOTA TEACHER ATTITUDE INVENTORY

by K. M. Evans

Faculty of Education, University College, Cardiff

SUMMARIES already published in this journal (Evans, 1958, 1959 and 1959a; Allen, 1962) have indicated the vast amount of research which has been carried out on the selection, assessment and personal qualities of teachers. Most of this work has been, at best, inconclusive and no clear picture has emerged of the types of people who are likely to be effective in the classroom. Success seems to depend quite as much on the school situation in which a teacher finds himself as on his own qualities.

It is fair to deduce that, given the right training, a wide variety of people can become at least average teachers, and attention is beginning to be turned to research on the training process. Inquiries like those by Warburton, Butcher and Forrest (1963) and Herbert and Turnbull (1963) may be cited in this connection.

Some of the tests used in these studies, such as the Allport-Vernon-Lindzey *Study of Values* and the Bernreuter *Personality Inventory*, are applicable to people in any walk of life, but there are a few which are specifically concerned with qualities of the teacher in the classroom. Among these last are a number of attitude scales, for example, those constructed by Yeager (1935), Evans (1946, 1952, 1953, 1965) and the better known Minnesota Teacher Attitude Inventory (Cook, Leeds and Callis, 1951).

Until recently, the Minnesota Teacher Attitude Inventory had not been much used in this country. Arvidson (1956) found it gave some indication of the efficiency of experienced teachers, but Evans (1958) reported that it appeared to be of little use as a predictor of the success of students in a one-year postgraduate teacher-training course. Then, in 1963, Herbert and Turnbull published a study of students in a three-year training course which included evidence that the *Inventory* distinguished between the more and less successful students. In view of its possible increasing use, it seems necessary to examine, in some detail, both the form and content of the M.T.A.I.

Description of test

The M.T.A.I. is designed to measure the attitudes of teachers towards their pupils and to distinguish those who are likely to establish a desirable classroom atmosphere from those less likely to do so. This involves definition of what constitutes a desirable atmosphere. In the opinion of the authors of the test, ‘ the teacher and pupils should work together in a social atmosphere of co-operative endeavour, of intense interest in the work of the day, and with a feeling of security growing from a permissive atmosphere of freedom to think, act, and speak one's mind with mutual respect for the feelings, rights and abilities of others ’. At the other end of the scale is the classroom where ‘ the teacher tends to think in terms of his status, the correctness of the position he takes on classroom matters, and the subject matter to be covered rather than in terms of what the pupil needs, feels, knows and can do ’. The assumption is made that the teacher's attitudes determine the type of classroom atmosphere he will maintain and that a liberal attitude towards his pupils is likely to characterize a good teacher. The test therefore seeks to pinpoint teachers' attitudes on a liberal-conservative continuum.

The M.T.A.I. consists of 150 statements of opinion about children and their behaviour, to each of which a response is required on a five-point scale: strongly agree, agree, undecided, disagree, strongly disagree. The statements are printed in a non-expendable test booklet while the responses are made on separate answer sheets.

For convenience, responses are designated ‘ Right ’ and ‘ Wrong ’, though it is recognized that these terms are incorrect and there are, in fact, no right or wrong answers, only agreement or disagreement with particular attitude statements.

The test is scored by using stencils giving the number of ‘ Rights ’ and ‘ Wrongs ’. The final score is: total ‘ Rights ’ *minus* total ‘ Wrongs ’. A positive score is taken as indicating a liberal (desirable) classroom attitude on the part of the teacher; a negative score suggests a repressive, dominating attitude.

Since there are 150 items—to each of which one of 5 responses must be made—there are in all $150 \times 5 = 750$ possible responses. Examination of the scoring stencils showed that 275 possible responses count as ‘ Right ’, while 326 count as ‘ Wrong ’. This leaves 149 which do not contribute to the score at all, and which may be called ‘ Neutral ’ for convenience.

At this point it may be useful to quote some of the statements in the test, and ten, selected by the use of a table of random numbers, are quoted in Table 1. For ease of reference, the ways in which the various responses are scored are shown in the column on the right, but it must be remembered that responses would actually be made on a separate sheet, not in the test booklet itself.

Table 1.

Sample Statements and Responses from M.T.A.I.

	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
3. Minor disciplinary situations should sometimes be turned into jokes	R	W	W	W	N
97. Teachers are responsible for knowing the home conditions of every one of their pupils.	R	R	W	W	N
16. A pupil's failure is seldom the fault of the teacher.	W	W	R	R	N
12. Pupils should be required to do more studying at home.	W	W	R	R	R
55. Aggressive children are the greatest problems.	W	W	N	R	N
84. A teacher should not tolerate use of slang expressions by his pupils.	W	W	W	R	R
63. Too much nonsense goes on in many classrooms these days.	W	W	R	N	R
33. The teacher should not be expected to manage a child if the latter's parents are unable to do so.	W	W	N	R	N
57. Many teachers are not severe enough in their dealings with pupils.	W	W	R	R	R
18. A teacher should never discuss sex problems with the pupils.	W	W	W	R	R

R =' Right '; W =' Wrong '; N =' Neutral '

Discussion of form of test

The fact that there is a large excess of ' Wrong ' over ' Right ' responses suggested that, although the possible scores range from +150 to −150, there is a much greater chance of obtaining a negative score than a positive one. To test this hypothesis, an answer sheet was filled in using a table of random numbers to determine the responses. The final score was +58 − 67 = − 9, supporting the hypothesis.

The next step was to ask a number of people to fill in answer sheets without reference to the test. This group included two lecturers, four students, a secretary, a housewife, a physiotherapist and a caretaker. They were given only answer sheets, unaccompanied by test booklets, and were asked to mark one response in each line, so that it was a purely mechanical operation and no question arose of their expressing any real opinions, since they were not presented with any statements about which to form opinions. Their scores are given in Table 2.

The most obvious characteristic of this group of scores is that they are all negative, again supporting the hypothesis that anyone who answers haphazardly is likely to get a negative score. This, in itself, is an undesirable feature. 'Guesses' should affect 'Right' and 'Wrong' scores equally if they are not to influence final scores, and this is usually the reason for using the 'Rights' *minus* 'Wrongs' formula rather than taking the 'Right' score as final. Here, the precaution is unavailing.

Table 2.

'Blind' Scores on M.T.A.I.

Subject	*'Right'*	*'Wrong'*	*'Neutral'*	*'Right' – 'Wrong'*
Lecturer 1	59	67	24	−8
Lecturer 2	50	54	46	−4
Student 1	58	80	12	−22
Student 2	57	70	23	−13
Student 3	40	82	28	−42
Student 4	51	73	26	−22
Secretary	53	67	30	−14
Physiotherapist	66	68	16	−2
Housewife	58	67	25	−9
Caretaker	53	74	23	−21
Average	54·5	70.2	25·3	−15·7
Random numbers	58	67	25	−9

Another type of careless answering was also considered. This involves the tendency to make a particular response irrespective of the content of the statement concerned. There are some people who tend to agree with any statement made, while others tend always to disagree, and yet others cannot make up their minds. How would these types fare on the M.T.A.I.?

To answer this question, an analysis was made of the numbers of 'Right', 'Wrong' and 'Neutral' responses under each of the five column headings. The results are given in Table 3.

Table 3.
Scoring Patterns of Five Types of Response

Response	*'Right'*	*'Wrong'*	*'Neutral'*	*'Right' – 'Wrong'*
Strongly agree	32	74	44	−42
Agree	25	116	9	−91
Undecided	52	62	36	−10
Disagree	88	52	10	+36
Strongly disagree	78	22	50	+56
Totals	275	326	149	−51

It is evident that stereotyped responses would produce meaningful scores. People who tend to agree or to be undecided are likely to be diagnosed as having a dominating classroom attitude, while those who tend to disagree are likely to be classed as establishing good rapport with pupils.

A test of this type should have built-in safeguards to prevent meaningful scores being obtained when it is filled in haphazardly, and it should also foil the person who makes a stereotyped response. People behaving in either of these ways are not really expressing any opinions, and their final scores should reflect this. This requires that there should be equal numbers of 'Right' and 'Wrong' responses available and that these, and the 'Neutral' responses, should be evenly distributed among the categories of response from *Strongly Agree* to *Strongly Disagree*. Insofar as it does not embody such safeguards, the form of the M.T.A.I. cannot be considered satisfactory.

Discussion of content of test

So far no account has been taken of the content of the test items, though examples have been given in Table 1, along with the ways in which the various responses are scored.

In the *Manual*, the authors say that the M.T.A.I. was validated by comparing the responses of two groups of teachers, one group designated by their principals as having excellent working relations with their pupils, the other as being unsatisfactory in this respect. The scoring patterns obtained in this way do not always appear to be logical. For example, in the case of No. 63 (*see* Table 1) *Undecided* and *Strongly Disagree* are both scored 'Right' while *Disagree* is classed as 'Neutral' and does not affect the score at all.

It was thought that it would be interesting to obtain M.T.A.I. scores from a series of groups at different educational stages, and to see whether any marked differences existed. Accordingly, the following groups were tested:

1. Forty sixth-form pupils in a boys' grammar school, who might be presumed to be relatively unsophisticated as far as the subject matter of the M.T.A.I. was concerned.

2. Two very small groups, consisting of second and third year university students taking a theoretical course in Education as one of two degree subjects.
3. Another small group of undergraduates, friends of those described in (2), not taking Education as a degree subject though some of them had taken it during their first year at the university.
4. A group of twenty volunteers from students who had just completed a one-year postgraduate teacher-training course.
5. Thirty experienced teachers interested in educational research.

The scores of these groups are given in Table 4.

Table 4.

Scores of Various Groups on M.T.A.I.

Group	*No.*	*'Right'*		*'Wrong'*		*'Neutral'*		*'Right' – 'Wrong'*	
		Mean	*Standard Deviation*	*Mean*	*Standard Deviation*	*Mean*	*Standard Deviation*	*Mean*	*Standard Deviation*
1. Sixth form boys	40	67·8	12·8	63·3	14·9	18·9	6·3	4·5	27·0
2. Undergraduates taking Education: 2nd Year	7	79·1	10·5	54·6	13·4	16·3	6·3	24·6	23·2
3rd Year	11	82·3	11·7	51·5	12·3	16·3	4·5	30·8	23·6
3. Undergraduates not taking Education	10	82·8	11·3	54·6	12·9	12·6	4·6	28·2	23·9
4. Postgraduate teacher-training students	20	83·2	14·3	54·2	14·5	12·7	3·4	29·0	28·6
5. Experienced teachers	30	76·8	15·7	58·7	16·9	14·4	5·8	18·1	33·4

It would be folly to attempt to draw any firm conclusions from scores from such small groups, but some observations may be made:

(1) The schoolboys made very much lower scores than either students or teachers.

(2) The experienced teachers made lower scores than any of the student groups.

(3) All the student groups scored at approximately the same level.

The results reported by Herbert and Turnbull (1963) are given for comparison.

Table 5.
M.T.A.I. scores reported by Herbert and Turnbull (1963)

Group	*Mean* (' *Rights* ' – ' *Wrongs* ')	*Standard Deviation*
1st year students	10.7	29.8
3rd year students	28.4	26.0

The first year students, fresh from school, did not score appreciably differently from the sixth form boys, while the third years produced scores in close agreement with those of the university students. This suggests that there is a considerable liberalization in views on teacher-pupil relations during student years, and this may be true whatever the subjects being studied. At the same time, it is noticeable that the standard deviations in all groups are high, showing the wide range of opinions held by individuals. More information on changes in attitudes in the immediate post-school years might be both interesting and useful, but the value of the M.T.A.I. is likely to be greatest in the field of teacher-training and it is more fruitful to consider it in this connection than generally.

The topics covered by the items of the M.T.A.I. are such as are commonly dealt with in teacher-training courses (*see* Table 1), and, in the main, the responses scored ' Right ' would be likely to be endorsed by many lecturers with liberal educational views. It is likely that their students will come to hold, or at least to express, such opinions. The more effective the lecturer and the more teachable the students, the greater the degree of indoctrination to be expected.

Both Evans (1958) and Herbert and Turnbull (1963) found a positive association between scores on the M.T.A.I. and marks in educational theory obtained by teacher-training students, while Herbert and Turnbull found also that the M.T.A.I. discriminated between students obtaining high and those obtaining low practical teaching assessments. The assessments of students' work, both practical and theoretical, are subjective and affected by the views about desirable pupil-teacher relations of the lecturers who make them. Students who adopt the views of a liberal lecturer are likely to be assessed favourably and also to score highly on the M.T.A.I.

This, however, is only a hypothesis, and it must be remembered that there was very little to choose between the mean M.T.A.I. scores of students taking Education or following a teacher-training course (Groups 2 and 4, Table 4) and those not doing so (Group 3). While Groups 2 and 4 might have been influenced by the views of Education tutors, Group 3 is unlikely to have been so influenced.

More research is needed to determine why students' M.T.A.I. scores increase, as they appear to do, during their college careers, and the attitudes of their tutors might be one factor to be investigated in this connection. The results of such a piece of work might have valuable practical applications in the field of teacher training.

The lower scores of the experienced teachers (Group 5, Table 4) as compared with those of the students may mean that their idealism has been tempered by classroom experience. Alternatively, it may mean that they no longer feel the necessity for reflecting the idealism of their tutors. No firm conclusion can be drawn about this without more information than is at present available.

Conclusions

It does not seem, from the available information, that the M.T.A.I. would be a satisfactory instrument for use in selection of students. Even though differences between mean scores of more and less successful students may be significant, the high standard deviations generally reported point to a considerable overlapping of scores in the two groups and consequent difficulty of discrimination among individual members. In any case if, as is possible, the M.T.A.I. is really measuring the extent to which students absorb the views of their tutors, a verbal intelligence test might give more satisfactory prognostications of success.

A final point should be noted. It seems quite definite that, if the M.T.A.I. is to be used in Britain, it will have to be re-standardized. All who have used it in this country agree that the American norms are not applicable here and it is also probable that some re-wording of items would be desirable. At the same time, the form of the test could be improved on the lines suggested earlier. Revised in these ways, the M.T.A.I. might provide more useful information than it does at present.

Bibliography

ALLEN, E. A. (1962). 'Professional training of teachers: a review of research'. *Educ. Res.*, V, 3, 200-215.

ARVIDSON, G. L. (1956). 'Some Factors influencing the Achievement of First Year Secondary Modern School Children'. Unpublished Ph.D. thesis, University of London Library.

COOK, W. W., LEEDS, C. H. and CALLIS, R. (1951). *The Minnesota Teacher Attitude Inventory.* New York: Psychol. Corp.

EVANS, K. M. (1946). 'A Study of Attitude towards Teaching as a Career'. Unpublished M.A. thesis, University of London Library.

EVANS, K. M. (1952). 'A study of attitude towards teaching as a career'. *Brit. J. Educ. Psychol.*, 22, 63-69.

EVANS, K. M. (1953). 'A further study of attitude towards teaching as a career'. *Brit. J. Educ. Psychol.*, 23, 58-63.

EVANS, K. M. (1958). 'An examination of the Minnesota Teacher Attitude Inventory'. *Brit. J. Educ. Psychol.*, 28, 253-257.

EVANS, K. M. (1959). 'Research on teaching ability'. *Educ. Res.*, I, 3, 22-36.

EVANS, K. M. (1959a). 'The teacher-pupil relationship'. *Educ. Res.*, II, 1, 3-8.

EVANS, K. M. (1965). *Attitudes and Interests in Education.* London: Routledge and Kegan Paul.

HERBERT, N. and TURNBULL, G. H. (1963). 'Personality factors and effective progress in teaching'. *Educ. Rev.*, 16, 24-31.

WARBURTON, F. W., BUTCHER, H. J. and FORREST, G. M. (1963). 'Predicting student performance in a university department of education'. *Brit. J. Educ. Psychol.*, 33, 68-79.

YEAGER, T. C. (1935). *An Analysis of Certain Traits of High School Seniors Interested in Teaching.* New York: Teachers College, Columbia University.

Research Note:

THE MEASUREMENT OF TEACHERS' ATTITUDES

by C. J. S. Tuppen

National Foundation for Educational Research

Introduction

THIS note gives details of six Guttman scales which have been designed for a study of attitudes of teachers in junior schools. Some information is also given on relationships between teachers' attitudes and other variables. A future publication will discuss the effects of teachers' attitudes upon their pupils. The attitude scales were constructed primarily for use in the N.F.E.R. Streaming Project, but other research workers may be interested to know of their existence.

Teachers' attitudes were studied as part of the Streaming Project, since it has been suggested that a teacher's effectiveness depends at least as much upon his attitudes as upon his length of experience or qualifications. This has never really been tested, but it seems more than likely that if a class teacher does not share the Head Teacher's beliefs about methods or organization or teaching, then these methods may not receive a fair trial, and the teacher himself may not be fully effective.

Research now in progress should give a clearer picture of the contribution of teacher attitudes to outcomes in streamed and non-streamed schools. Almost a thousand teachers are involved in the study; half in streamed and half in non-streamed schools.

Interviews and questionnaire

This present study was begun by carrying out a total of thirty-one unstructured interviews with junior school teachers. The interviews were divided among three interviewers, who acted as non-critical listeners and who tried unobtrusively to develop various themes which they had been briefed to investigate. Tape recordings were made and transcribed later. The interviews yielded a wealth of material which has been summarized elsewhere (Tuppen, 1965).

The interviews suggested a number of attitude-areas which appeared suitable for construction of attitude scales. These scales would make it possible to assess teachers' attitudes by means of a questionnaire. Such a device would of course not produce the detail obtainable from individual interviews, but would have the advantage that information could be obtained from a large number of teachers, taking up relatively little of their time, using a standard form for all teachers which would elimate interviewer bias. An incidental finding from the interviews was that teachers are not necessarily aware of the attitudes of other members of staff, so it would have been unwise to rely on, for example, a Head Teacher's assessment of the opinions of his staff.

A questionnaire was constructed, containing statements which had actually been made by teachers in the interviews. In the questionnaire instructions, each teacher was asked to indicate his degree of agreement or disagreement with each statement, using a five-point scale with a choice of responses ranging from 'strongly agree' to 'strongly disagree'. It is important to notice that the questionnaire was based upon the concepts of practising teachers and was phrased in their language; it was not drawn up in accordance with any theory of personality, nor were the items drawn from the educational literature. The use of preliminary exploratory research in the schools is the keystone of this method.

The scaling of items

The questionnaires were used to assess each teacher's attitudes relative to those of other teachers. Guttman's method of scalogram analysis was used for selecting items and for testing the assumption that there exists a unidimensional scale of (e.g.) 'attitudes towards streaming', one end of which could be labelled 'pro-streaming' and the other 'anti-streaming'. An individual's probable position on the scale could be inferred by studying the way in which he replied to a series of statements about streaming, provided that the items yielded mutually consistent responses. For example, two items from the scale were: (1) 'With non-streaming, I would find it impossible to keep duller children occupied while bright children recieved instruction', and (2) 'In a streamed school, one gets far more done for the slow learner'. The items could be used to distinguish three shades of opinion about streaming, as shown by the three response patterns (A, B, C) below:-

	Item 1.	*Item* 2.
A	Agree	Agree
B	Disagree (or undecided)	Agree
C	Disagree (or undecided)	Disagree (or undecided)

These items are said to be 'scalable', since they elicit mutually consistent response patterns. Everyone who agrees to item (1) also agrees to item (2). It is unlikely that anyone will agree to (1) and disagree with (2).* Pattern A is typical of a person at the pro-streaming end of the scale and was given by 11% of the teachers involved in the streaming project; pattern B is slightly less pro-streaming and was given by 39% of the teachers; pattern C is typical of someone nearer the anti-streaming end and was given by 50% of the teachers. (Note that streamed and non-streamed schools were equally represented in the sample.)

Ideally, a Guttman scale should contain about ten items. This is much shorter than a Likert or Thurstone scale. In practice, it is seldom possible to obtain as many as ten suitable items. Perfect consistency in responses is never achieved, but an attitude area is considered to be scalable if the proportion of inconsistent responses does not exceed 10%. Various other checks are also applied. In some attitude-areas, people's views are made up of so many elements that response consistency cannot be achieved and this method of attitude measurement cannot be used. This occurred in the author's attempt to produce a scale of 'attitudes towards backward children'.

---

*Though, of course, the reverse does not necessarily hold.

The attitude scales

Six attitude scales were however successfully constructed. They have been named as follows:

Scale A — 'Permissiveness'*
Scale B — Attitudes towards physical punishment
Scale D — Attitudes towards eleven plus selection
Scale G — Attitudes towards noise in the classroom
Scale H — Attitudes towards streaming
Scale I — Attitudes towards children in A streams

Co-efficients of reproducibility for these scales were all above .90, the values being obtained from a random sample of 100 teachers and replicated with a second random sample of the same size. The scales all contained at least five items, some of which were trichotomous.

The scales were found to be positively inter-correlated, indicating the existence of a general factor. Product-moment correlations are given below.

	A	B	D	G	H	I
A. Permissiveness	—					
B. Physical punishment	·29	—				
D. Eleven plus selection	·22	·24	—			
G. Noise in the classroom	·31	·25	·31	—		
H. Streaming	·23	·27	·39	·34	—	
I. A streams	·19	·16	·32	·27	·54	—

Comment

There was also some relationship with the types of lesson preferred by the teacher. A comparison was made in the *Non-streamed* schools of teachers holding extreme attitudes for and against streaming; this revealed that the pro-streamers made relatively more use of 'traditional lessons' such as formal sums, learning lists of spellings, while they made less use of 'progressive lessons' such as projects and free activity.

In general, teachers in non-streamed schools were found to be more 'progressive' on all six attitude scales than teachers in streamed schools. They also tended to be younger, and younger teachers tended to hold more 'progressive' opinions—particularly in their rejection of eleven plus selection, their permissiveness, and their tolerance of noise. (These differences were significant at the 1% level.) Thus streaming and non-streaming are not merely different methods of organization—teachers do hold different attitudes and schools do embody different philosophies about teaching. But it would be quite wrong to suppose that all teachers favoured the type of organization used by their school. The table below makes use of the 'attitudes to streaming' scale to show the percentages of teachers favouring streaming, holding neutral attitudes, or favouring non-streaming.

---

*This scale attempted to assess teacher attitudes in terms of the type of classroom behaviour which they elicited or permitted, as revealed by their responses to five statements.

The statements were: 'Opportunities for self-expression through movement, painting and writing poetry, are more important than concentrating on the "three R's"'. 'Naturalness is more important than good manners in Juniors'. 'Children must be taught to have decent manners'. 'Teachers should demand clean hands in school'. 'I cannot stand children fidgeting in class'.

	Pro-streaming	*Neutral*	*Anti-streaming*	*Total*
Teachers in streamed schools . .	45%	45%	10%	100% (N=495)
Teachers in non-streamed schools . .	13%	35%	52%	100% (N=473)

It is to be expected on the basis of this table that one teacher in any school which had about ten classes would definitely prefer the alternative type of organization. This expectation is rather misleading, for in a few schools it was found that less than half the staff were favourable to the school's policy, while in other schools all were in agreement.

Acknowledgements

The co-operation of the teachers who were involved in this study is gratefully acknowledged; as is the help of my colleagues, Dr. J. H. G. Hewlett, Mrs. J. Barker Lunn and other members of the staff of the N.F.E.R.; also the advice which I have received from Mr. A. Lunn on scalogram analysis.

References

STOUFFER, S. A., GUTTMAN, L., *et al.* (1950). *Measurement and Prediction.* Princeton, N.J.: Princeton University Press.

TUPPEN, C. J. S. (1965). 'A study of the expressed attitudes of junior school teachers towards aspects of teaching, using a Guttman scaling method'. Unpublished M.A. thesis, University of London.

Research Note:

SELECTED FINDINGS FROM A FIVE-YEAR STUDY OF ACADEMIC DETERIORATORS IN A GRAMMAR SCHOOL*

by R. R. Dale

Department of Education, University College of Swansea

and S. Griffith

Introduction

THIS article had its origin in the concern of one of the authors about those children who entered the grammar school in a high position on the entrance examination list and yet had a very low standard of attainment at the end of the first academic year. The authors therefore made a detailed study of all cases of marked deterioration which arose in one particular grammar school during a period of five years. The pupils were selected by a fixed and severe criterion of demotion, defined as follows.

Pupils on entry were streamed according to order of merit in the entrance examination. Demotion occurred at the end of both the first term and the first year. In the first four years of the inquiry there was a three-stream entry, and ' deteriorators ' were defined as pupils who were demoted by the end of their first year from an ' A ' to a ' C ' stream, or who were demoted from the ' B ' stream to the ' C ' stream and then fell to the lower half of it. Those who fell from an ' A ' to a ' B ' stream were excluded because the investigators wished to concentrate on what they regarded as the more serious cases. In the fifth year of the inquiry, when there was a four-form entry, the definition was adjusted accordingly. We accepted the limitations of this definition because we wished to examine the causes of deterioration rather than to determine its incidence.

The number of deteriorators isolated in this way was thirty-nine; though this is acknowledged to be a small number, the smallness was compensated by the intimate knowledge that we had of each deteriorator. One of the authors was a teacher in the school, in daily contact with the deteriorators usually for several years, who knew their home conditions well.

During the interval between the commencement of the research and the writing of the present article several large sociological surveys have demonstrated the importance of social class factors in connection with attainment at the grammar school. Such reports were of necessity impersonal, and here in a grammar school was flesh for the bones of these

---

*A fuller account of this study may be read in the recently published book: *Down Stream: Failure in the Grammar School* by R. R. Dale and S. Griffith. Routledge and Kegan Paul, 1965.

Acknowledgements:

We are grateful to the Headmaster of the School and the County authorities for their helpful co-operation over a long period, and to Dr. Phillip Williams for kindly reading the manuscript and offering some useful suggestions. The Department of Education and Science is thanked for providing clerical assistance.

reports. We were able to observe the pupils in their day to day conduct. We studied them in their school activities and went to see them in their leisure moments at home. We think that we came to know them. In some cases the inquiry leads back to the Primary School, in others to an examination of academic factors in the Grammar School, in others to the homes of the children. Always in the centre of the picture is the child; John, who is failing in spite of intellectual promise; Mary, who is succeeding in spite of her parents; David, dull and industrious; Elaine, bright and lazy. In the centre is the child, but on each side of him is a parent and in the background Society. Much detailed investigation therefore lies behind the statistics presented here.

The co-educational grammar school where these studies were made is situated in a small industrial area with an agricultural background. The percentage intake was very high, varying from 39 per cent. of the relevant age group to about 46 per cent. As, however, our criterion of deterioration excluded from consideration those pupils who were in the last thirty in the extrance examination, this research is about pupils who either come within the normal range of ability for grammar school intake or are reasonably near to it. As the marks in the 11-plus examination are close together at the border-line, and selection here is somewhat chancy, there is little difference between those who are just above and those who are just below the line.

Each year, after the demotions had been decided, each ‘ deteriorator ’ was tested and interviewed. They were given Schonell tests in word recognition and diagnostic arithmetic, and intelligence tests. The results of several other intelligence tests were available for a control group. Primary school records were scrutinized, and absenteeism and ill-health noted. The Principal School Medical Officer for the County supplied the health grading of each deteriorator. Inquiries were made about their home background and their behaviour outside school. Favourite and disliked subjects were listed, together with hobbies and other non-school activities. In cases of known misdemeanour the Probation Officer was consulted, and her knowledge sought about all the other deteriorators. Primary school staff, Youth leaders and Ministers of religion were also interviewed. As soon as possible the home of each deteriorator was visited at least once. The career of each of these pupils was followed until they left school. Rather late in the investigation it was decided, for the sake of contrast, to include ‘ improvers ’ in the study. These were selected by an objective promotion criterion similar to that used for demotion, though of course in reverse. Data for the improvers was necessarily more limited than that for the deteriorators. The work presented here is only a small part of the investigation. We have, however, selected those findings which seem to be of most importance. One of the first variables which we examined was ‘ intelligence ’ or the ‘ academic aptitude ’ of the pupils.

Intelligence

At the end of their first academic year group intelligence tests were administered to all the deteriorators. For a control group of 88 pupils, being a cross-section from all streams, use was made of the intelligence test taken in the entrance examination. The average I.Q. of the deteriorators was 111.4, and that for the control group 110.3. It is reasonable to assume from these figures that the deteriorators as a group were not handicapped by being of lower intelligence than their fellow pupils.

The I.Q. range of the deteriorators was as follows:

Table 1

I.Q.	*Below* 100	100–104	105–109	110–114	115–119	120–124	125–129	130 +
No. of Deteriorators	—	2	15	13	6	3	—	—

Average I.Q. +111.4

At first sight it might seem that the intelligence level of some of these deteriorators was rather low for grammar school work and indeed this factor may have played some part in the deterioration. In order to throw further light on this problem we compared the I.Qs. of deteriorators with those of the non-deteriorators of the same sex who were in corresponding positions in the same entrance examination. We were able to do this for Years 4 and 5 of the survey, comprising only 14 of the deteriorators, but we found that whereas the average I.Q. of the deteriorators was 112.9, that of the non-deteriorators was only 110.2. We see therefore that pupils who had the same or slightly lower I.Qs. than the deteriorators did not themselves become deteriorators. This indicates, as one would expect, that factors other than intelligence have a strong influence on academic success. In the clinical case study work we came across only two pupils where we concluded that lack of ability was the principal factor associated with deterioration though there were two other cases where it was a strong contributory factor.

Social class

The association of social class groupings with the standard of attainment within the grammar school became noticeable during the inquiry. This factor was also associated with the deterioration of the pupil while at the Grammar School. The table below presents a classification of parental occupation for:

(*a*) one whole year of entry [to serve as a base-line for interpreting the figures given at (*b*) and (*c*)];

(*b*) the 'improvers' for Years 1 to 6 inclusive;

(*c*) deteriorators for Years 1 to 5 inclusive.

The improvers included all pupils who conformed to a promotion criterion; they are presented here because they bring the social class factor into even greater prominence. It should be noted that the one year of entry will in itself contain some deteriorators and some improvers.

Table 2

Classification for Social Status from Parental Occupation

Parental Occupation Category:	1	2	3	4	5	6	7
(*a*) No. of pupils—one year	6	9	10	9	41	21	6
(*b*) No. of improvers—Years 1 to 6 inclusive	0	6	7	8	15	0	0
(*c*) No. of deteriorators—Years 1 to 5 inclusive	0	0	0	2	15	12	10

The categories of occupation in Table 3 are as follows:

1. Professional and high administrative.
2. Managerial and executive.
3. Inspectional, supervisory and other non-manual.
4. As for 3 but lower grade.
5. Skilled manual and routine grades of non-manual.
6. Semi-skilled manual.
7. Unskilled manual.

In the table the deteriorators tend to congregate in social class categories 5, 6 and 7, i.e. 37 out of 39 are in these categories; on the other hand the improvers do not go below category 5*. The fathers of 22 of the 39 deteriorators were unskilled or semi-skilled, i.e. 56.4 per cent, whereas only 26.5 per cent of the pupils in a whole year's entry (including future deteriorators) belonged to these classes. This finding is confirmed by other researches, notably the Ministry of Education report, *Early Leaving*, page 18, where it is stated that:

‘ Deterioration which has caused many who were placed in the top selection group at eleven, to be found at sixteen in the lowest academic categories, is most common among the children of unskilled workers (54 per cent) and semi-skilled workers (37.9 per cent) '.

---

*Statistical comparisons (Chi-square) of the social class distribution of deteriorators compared with the one-year entry (itself ‘ contaminated ’ with improvers and deteriorators) gives a significant difference well beyond the .025 level. The comparison between deteriorators and improvers gave a highly significant difference. That between improvers and the one-year entry was significant beyond the .05 level. Yates' correction was used in each case.

At the other end of the table none of the 39 cases of deterioration gathered from five years of entry, came from social classes 1, 2 and 3, whereas 13 out of the 36 improvers came from these classes. Even in the middle of the scale class 4 contains 8 out of the 36 improvers but only two of the 39 deteriorators.

In an attempt to isolate more specifically the particular influences associated with deterioration we included a detailed examination of a number of factors in the social class ' constellation '. We have selected our findings on two of the factors for report here. The first of these is the education of the parents.

Parental education

We began with a control group consisting of the total entry for two years and tabulated the pupils according to stream and education of parents. This information is given in Table 3.

Table 3

Streaming in relation to Education of Parents (percentages)*

	' *A* ' *stream*	' *B* ' *stream*	' *C* ' *stream*	
Number of children (Years 7 and 8)	64	63	54	
(*a*) No parent attended grammar school	15.1	31.5	53.4	100%
(*b*) Only mother attended grammar school	32.3	38.7	29.0	100%
(*c*) Only father attended grammar school	51.2	39.5	9.3	100%
(*d*) Both parents attended grammar school	61.8	32.4	5.9	100%

There were 108 families out of 181 with one or more parents who were taught at a grammar school. It should be noted in passing that whereas 83 per cent of children in the ' A ' stream had at least one parent who was educated in a grammar school, this proportion declined to 63 per cent in the ' B ' form and only 28 per cent in the ' C ' form. [Cf. Whalley (1961) pp. 138-140, and pp. 45-57]. From these figures alone one cannot say that the differences are produced solely by a decline in parental encouragement, as such factors as the correlation between social class and academic aptitude are also involved. The case study evidence, however, on progress in the grammar school, points strongly to parental influence as a major factor. The differences are most pronounced when both parents attended a grammar school, and when only the father attended a grammar school. When only the mother attended a grammar school the differences, as shown in the Table.

---

*The difference between (*a*) and (*b*) is significant beyond the .05 level, and those between (*a*) and (*c*) and (*a*) and (*d*) beyond the .01 level.

do not appear as great, but one would not care to make a definite finding from a single instance and with such small numbers.

We then examined the parental education of the deteriorators and found to our surprise that in *only one case amongst the* 39 *was there a parent educated in a grammar school.* whereas we have seen that in the control sample of a one-year entry there were 108 out of 181 families with one or more parents who were taught at a grammar school*. The research of Jones (1962) gives general support to these findings. His research produced a statistically significant association between deterioration and neither of the parents having been educated at a grammar school, and between deterioration and the mother not having been educated at a grammar school, but the difference caused by the father not receiving a grammar school education was not statistically significant, though the trend was the same. Further support for the influence of the mother on the educational progress of the child comes from our examination of the parental education of the full 36 improvers discovered in the entries of Years 1 to 6. Both the parents of one third of the improvers were educated at a grammar school, only the mothers of another third (30.6%) were so educated, and in the case of 19.4 per cent of improvers only the father was so educated. In only 16.6 per cent of the cases did neither parent attend a grammar school. We are supported in this finding by the work of Greenald (1955), who concluded that there was some association between improvement and the parents' further or grammar school education†.

The scarcity of grammar-school parents amongst the deteriorators does not, of course, imply an inevitable lack of academic culture in the home, but it does mean that all these parents have missed the experience of grammar school life and studies which could help them with their children's difficulties; and it cannot be easy for them to understand fully what the school requires of them as parents. Lewis (1952) writes '. . . it was often more difficult for a boy from a poorer home to "fit in" as naturally as a boy whose parents had been to grammar schools, and where their own backgrounds and interests were not unlike those of the schoolmasters the boy would meet in school'.

The second factor selected was size of family.

Family size

While we were examining the association between family poverty and deterioration in attainment we came across important differences concerning family size. This is an aspect of the question which merits much more attention. As Stott (1956) states: 'When one examines the home background of delinquents or even of children who are backward at school, one cannot help remarking how frequently they are members of large families, and recent studies confirm that the numbers coming from such families are indeed disproportionate'. Even when social class is held constant, this effect of family size can be seen at work. (Cf. Floud, *et al.*, Table 21.)

Not only is a large family more likely to draw near the poverty line, and, therefore, be unable to pay for facilities such as books, but also, interruption by other children in the family may help substantially towards the deterioration of a grammar school pupil. Four of our deteriorators suffered badly from such interruptions. Nor can the parents give as much attention to any one of five or six children as they could to an only child.

---

*The difference is statistically highly significant.

†Cf. also the important book by Fraser, E. (1959), pp. 41-43.

The difference between the size of the families of the deteriorators and those of a control group is shown in Table 5.

Table 4

Size of Family

	Number of children in family					
	6 *or more*	5	4	3	2	1
	%	%	%	%	%	%
One-year entry (95 pupils not counting deteriorators)	—	8.7	21.7	17.4	39.1	13.0
Deteriorators in the survey (39) (Years 1 to 5)	18.0	5.1	28.2	15.4	20.5	12.8

This table was prepared by noting the number of children in the family of each deteriorator, towards the end of the pupil's first year of attendance.

It will be seen that whereas 51 per cent of the deteriorators belonged to families where there were four or more children, only 30 per cent of the non-deteriorators were similarly placed,* and in the entry year considered in Table 4 there was no family of a non-deteriorator which had more than five children. The average size of the families of deteriorators was 3.9 whereas that of the 36 improvers was 2.1. Greenald (1955) came to the same conclusion. He reported that a family size of four children or more was associated with a downward trend in academic performance, in both middle and working classes. Academic improvement was associated with only children, particularly among the working class. In this study 67 per cent of the improvers were the first-born of a small family; the figure for deteriorators was 48.7 per cent.*

In order to ascertain more clearly the inter-relationship between social class, size of family and pupil attainment the data was reorganized as it appears in Table 5.

---

*The differences are statistically significant.

Table 5

Size of Family and Social Class

Deteriorators compared with Improvers*

Children in Family		*Social Class* 2	3	4	5	6	7
1	Deteriorators				2	2	1
	Improvers	(1)	(4)	(3)	(5)		
2	Deteriorators			1	1	3	3
	Improvers		(3)	(3)	(7)		
3	Deteriorators				3	2	1
	Improvers	(4)		(2)	(1)		
4	Deteriorators			1	4	2	4
	Improvers				(2)		
5	Deteriorators				1	1	
	Improvers						
6	Deteriorators				2	1	
	Improvers						
More than 6	Deteriorators				2	1	1
	Improvers	(1)					

The dotted line is drawn to include all but one of the 36 improvers. When one reads from left to right, there are no improvers below social class 5 and when one reads downwards there is only one improver from families of more than 4 children. In this exceptional case the father was a professional man (social class 2). Within each social class the family size of the improvers is smaller than that of the deteriorators; in social class 4 it is 1.9 compared with 3.0 and in social class 5 it is 2.0 compared with 4.9. The other classes do not overlap, but for social class 3 the average size for improvers was 1.4, and in social class 2 it was 3.6 (because of the exceptional family of 9); for deteriorators in class 6 the average size was 3.6 and for those in class 7, 3.4. As the number of deteriorators was only 39, and there were only 36 improvers, we record this information mainly to enable comparisons to be made with the results of other investigations. Once again, however, we feel it necessary to remind readers of the ' floor ' and ' ceiling ' effects due to our definition of improvers and deteriorators. A large proportion of the higher social class children were in the ' A ' stream on entry, because of their high place in the entrance examination. They were therefore more ' at risk ' than those from the lower social classes except that they were not included as deteriorators if they went down only to the ' B ' stream. Pupils in the ' A ' stream could not, however, become improvers. On the other hand there was a large proportion of pupils of low social class in the ' C ' stream. It was therefore statistically easier for them to produce improvers, whereas they could not, by definition, become deteriorators. Our working definitions, therefore, tend to produce an underestimate of the association between social class and attainment in the grammar school.

---

*Improvers are in brackets; they are from Years 1 to 6, whereas deteriorators are from Years 1 to 5.

References

FLOUD, J. E., HALSEY, A. H. and MARTIN, F. M. (1957). *Social Class and Educational Opportunity.* London: Heinemann.

FRASER, E. (1959). *Home Environment and the School.* London: U.L.P., for the Scottish Council for Research in Education.

GREENALD, G. M. (1955). ' An inquiry into the influence of sociological and psychological factors on the trend of achievement in grammar school pupils '. M.A. thesis, London University.

JONES, C. V. (1962). ' An inquiry into the causes of gross discrepancy between the performance of pupils in the 11 plus examination and their performance at the end of the first year in a Welsh grammar school.' M.A. thesis, University College of Swansea.

LEWIS, B. R. (1952). ' Some relationships between comparative success and failure in Grammar Schools and certain aspects of personality '. M.A. thesis, London University.

MINISTRY OF EDUCATION Central Advisory Council for Education (1954). *Early Leaving.* London: H.M.S.O.

STOTT, D. H. (1956). *Unsettled children and their families.* London: U.L.P.

EXPERIMENTAL EVIDENCE ON COMPARATIVE ATTAINMENT IN MATHEMATICS IN SINGLE-SEX AND CO-EDUCATIONAL SECONDARY SCHOOLS

by W. H. King

University of London Institute of Education

Introduction

IN two previous articles in this journal, R. R. Dale has considered the available research evidence on attainment in Mathematics and other subjects of boys and girls in secondary schools. Several difficulties are inherent in any attempt to generalize or to give a valid interpretation of these results. There are many variables: social class, rural *v.* urban schools, degree of motivation of students and parents; these have rarely been considered. Most of the research reported by Dale is based on raw examination marks or passes in the old School Certificate or present G.C.E. ' O ' level examinations. King's (1949) research dealt with a random sample of School Certificate results abstracted from the examination broadsheets, but, in common with other researches, did not allow for:

(i) number of students who did not take the examination (the 'drop outs ');

(ii) pass or fail in a public examination as a statistically rough criterion; or

(iii) adequate comparison of attainment of boys and girls in co-educational schools.

Items (i) and (ii) may give misleading results, and these difficulties may be avoided if complete forms or year groups can be tested before:

(a) the compulsory school leaving age has been reached, and

(b) any of the students have dropped the subject.

In addition, results must be based on attainment of children in randomly selected secondary schools. Such information was available from a research reported in this journal by King (Vol. VIII., No. I.).

Sampling

A stratified sample of 46 secondary schools gave the following breakdown into types of schools. The numbers of schools and students who completed all tests are also recorded:

Table I

Type of School	*Sex*	*Single Sex*		*Co-educational*		*Total Students*
		No. of Schools	*No. of Students*	*No. of Schools*	*No. of Students*	
Grammar	Boys	7	417	5	167	584
Grammar	Girls	7	406	5	153	559
Modern	Boys	7	367	13	354	721
Modern	Girls	7	345	13	347	692
	Boys		784		521	1305
	Girls		751		500	1251
	Total	28	1535	18	1021	2556

Results

In this report, results are given for three variables:
Scores on:

(a) Raven's Progressive Matrices (1956);
(b) N.F.E.R. Mathematics Test I;
(c) Step 3A Mathematics Test;

for boys and girls separately in co-educational schools and for boys and girls in single-sex schools.

Table II summarizes the results of boys and girls on these three tests.

Table II

Test	*Sex*	*Mean*	*N*	*Significance*
Raven's Progressive Matrices	Boys	46.35	1305	.05*
	Girls	47.09	1251	
N.F.E.R. Mathematics	Boys	68.98	1305	N.S.*
	Girls	68.37	1251	
Step 3A	Boys	31.29	1305	.01*
	Girls	29.55	1251	

*N.S. The difference between the two means is not significant, e.g. could have occurred by chance alone.
.01 A difference between the two means as large as this may be expected by chance only once in every hundred occasions.

In these results, the girls' scores are significantly higher than the boys' scores on the non-verbal test, not significantly different on the standard N.F.E.R. Mathematics test. The Step 3A is an American mathematics test quite different in content and presentation from the usual English test. Any rigidity in mathematical thinking would tend to show here by lower scores; this may have been the reason that the boys' scores are very significantly higher than the girls' scores. However, the break-down of these results for boys and girls in co-educational and single-sex schools revealed further variations in means. These results are presented for each of the three variables.

Variable: Raven's Progressive Matrices

(a) Grammar School Sample

Calculation of mean scores for boys and girls provided range of mean scores in order of size.

	Girls		*Co-ed: Girls*		*Co-ed: Boys*		*Boys*
Mean Scores	52.1		50.9		50.6		49.4
Significance		.02		N.S.		.01	

All other possible mean differences are highly significant (.01 level).

(b) Modern School Sample

	Boys	*Girls*	*Co-ed: Girls*	*Co-ed: Boys*
Mean Scores	46.9	44.3	42.3	40.2

Here there is a very significant difference (.01) between any two means though there was no signicant difference between the means of all the boys and all the girls in this sample of secondary modern students.

Variable: N.F.E.R. Mathematics Test I

(a) Grammar School Sample

The differences in boys' and girls' means in single-sex schools and boys' and girls' means in co-educational schools were not significant. The breakdown gave the following range of mean scores:

	Girls		*Co-ed: Boys*		*Boys*		*Co-ed: Girls*
Mean Scores	87.19		86.87		86.80		84.71
Significance		N.S.		N.S.		.05	

Only two differences reached significance level; girls' from single-sex schools and girls' from coeducational schools had means which differed very significantly (.01) and boys' means from single-sex schools were higher (.05) than those of girls from co-educational schools. All other differences were not significant.

(b) Modern School Sample

There was no significant difference between the mean scores of boys (721) and girls (692) in this sample, but the mean scores under the four categories do show great variations.

	Boys		*Girls*		*Co-ed: Girls*		*Co-ed: Boys*
Mean Scores	62.09		60.23		47.24		46.68
Significance		N.S.		.01		N.S.	

All other mean differences are highly significant (.01). Attainment of boys and girls in separate schools is significantly higher than that of boys and girls in co-educational schools on this conventional type of test. There was a wide range of mean scores for children in the co-educational schools though such differences were not as large in the other two more unusual types of tests. The range of means for girls was 23.0 to 75.6 and for boys 24.0 to 82.1., both from co-educational schools. With the omission of the two coeducational schools with smallest means, the boys' mean rose to 49.8 and the girls' to 50.7 but were still significantly lower than the means for single-sex schools. The wide range of mean scores has already been noted*.

Variable: Step 3A Mathematics Test

(a) Grammar School Sample

The mean scores for all boys and all girls were respectively 39.11 and 38.28, a difference which, though small, was significant (.02). The breakdown into the four categories is made as before.

	Boys		*Co-ed: Boys*		*Girls*		*Co-ed: Girls*
Mean Scores	39·13		39·04		38.70		37·14
Significance		N.S.		N.S.		.01	

Here, the only differences to note are those between the mean of girls in coeducational schools and the mean of each of the other three categories (.01), otherwise most students make the same type of response to this unconventional test.

(b) Modern School Sample

The mean scores of all the boys was significantly higher than the mean scores of the girls (.01)—a larger difference than had appeared in the grammar school sample. The breakdown into different categories produced more consistent differences than any previously noted.

	Boys		*Girls*		*Co-ed: Boys*		*Co-ed: Girls*
Mean Scores	27·34		23.82		22.50		21.20
Significance		.01		.05		.05	

Boys in single-sex schools did much better (.01) than the students in the other three categories; girls better than girls in co-educational schools (.01). This test produced the largest variation in means.

Discussion

Overall comparison of boys' and girls' attainments in Mathematics is meaningless because such comparison disguises differences between the groups when considered in terms of types of school and the nature of the criteria to be measured. Dale, in his discussion of previous research, states that, in general, boys from co-educational schools do better than boys from boys' schools, but there is less difference between girls in co-educational schools and girls' schools. The first statement has been verified only for the grammar school boys, but not for the secondary modern boys. The second statement finds no support in either sample.

Many of the differences between grammar and modern sample means must be due to secondary selection and its efficiency in selecting boys of similar academic quality, in addition to differences demanded in academic standards and degrees of motivation.

There is some agreement of these results with Sutherland's Ulster research. She found that boys from co-educational schools had higher mean scores on both types of mathematics paper in the Senior Leaving Certificate Examination than boys from boys' schools. This pattern of differences was repeated on the N.F.E.R. test, but not on the Step 3A, though in both cases the differences were not significant.

The differences in means for both types of girls' schools depended on the particular mathematics test; but in the present research, girls from girls' schools obtained significantly higher scores on both tests. Girls obtain higher mean scores on the non-verbal test in the grammar school sample and do significantly better than boys from co-educational schools in the modern sample.

It is impossible to give a brief summary, but two points are clear:

(a) the superiority of girls on the non-verbal test and of boys on the Step 3A test in the grammar school sample, and

(b) boys and girls from single-sex modern schools do consistently better than those from co-educational schools.

Bibliography

DALE, R. R. (1962). 'Coeducation: a critical analysis of research on the effects of co-education on academic attainment in grammar schools'. *Educational Research*, IV, 3.

DALE, R. R. (1962). 'Coeducation, II: an analysis of research on comparative attainment in mathematics in single-sex and co-educational maintained grammar schools'. *Educational Research*, V, 1.

KING, W. H. (1949). 'A critical analysis of the results of a school certificate examination'. Unpublished Ph.D. thesis, University of London.

KING, W. H. (1965). 'Differences in mathematical achievement related to types of secondary schools and their geographical locations'. *Educational Research*, VIII, 1.

SUTHERLAND, M. (1961). 'Coeducation and school attainment'. *Brit. J. Educ. Psych.*, 31, 2,

VOLUME VIII, No. 3 June, 1966

EDUCATIONAL RESEARCH

A Review for Teachers and all concerned with Progress in Education

Published for
THE NATIONAL FOUNDATION FOR EDUCATIONAL
RESEARCH IN ENGLAND AND WALES
by
NEWNES EDUCATIONAL PUBLISHING CO. LTD.
Tower House, Southampton Street, London, W.C.2

© National Foundation for Educational Research in England and Wales
Printed by King, Thorne & Stace Limited, School Road, Hove, Sussex
and published by Newnes Educational Publishing Co., Ltd., Tower House, Southampton Street, London, W.C.2.,
for the National Foundation for Educational Research in England and Wales, The Mere, Upton Park, Slough, Bucks.

COLOURED IMMIGRANT CHILDREN: A SURVEY OF RESEARCH, STUDIES AND LITERATURE ON THEIR EDUCATIONAL PROBLEMS AND POTENTIAL — IN BRITAIN*

by R. J. GOLDMAN and FRANCINE M. TAYLOR

Education Research Unit, University of Reading

I. General background of the racial situation in Britain

THE situation of coloured immigrant children in our schools today can only be fully understood against the background of the total immigration situation, and the way in which it has changed in recent years. Our introduction, therefore, outlines the general situation and the studies stemming from it.

Commonwealth coloured immigration has become a controversial issue in Britain today. Although the total estimated number of immigrants at the end of 1964 was given as 820,000 by E. J. Rose, Head of the Five-year Study of Race Relations in Great Britain, and by the Home Secretary in the House of Commons (*The Times*, 18 January, 1965), this is considered by some to be an underestimate, in view of the uncertainty expressed in the Census Report, 1961, and that the true figure is closer to a million. Although 820,000 is only 1·5% of the total population of 54 millions, over 80% of coloured immigrants are, according to Peach (1965), concentrated in the urban areas (towns of 50,000 or more), with half in Britain's ten largest cities. This means that in some towns and cities the concentrations of coloured immigrants is as much as 8% and in particular neighbourhoods as much as 50% or more.

Coloured immigrants are drawn mainly from three ethnic groups whose approximate totals probably include 430,000 West Indians; 165,000 Indians; 100,000 Pakistanis; and in addition, some 125,000 come from African countries, or from Aden, Ceylon or Malaya. There are other large groups of Commonwealth citizens thought to number about 60,000 Cypriots, 38,000 Maltese, 60,000 Australians and New Zealanders, and 60,000 Canadians. The largest immigrant group comes from the Republic of Ireland, an estimated 500,000 (B.B.C., 1965).

The main reasons behind immigration to the United Kingdom appear to have been poverty, lack of economic and educational opportunity in the home country, the availability of jobs in Great Britain and the active recruitment of foreign workers by some industries in this country. The main arrival of Commonwealth citizens began in the early 1950s and rose from an estimated 42,000 in 1955 to an estimated 136,000 by 1961 [Home Office (Cmd. 2739), 1965]. The number of immigrants who arrived immediately before the Commonwealth Immigration Act (1962) rose sharply in anticipation of the controls, whereby the Act regulated the entry of those who wished to work by determining the number of Ministry of Labour vouchers allowed. Under the first Act approximately 20,000 vouchers a year were issued: 'A' type vouchers to those with specific jobs to come to; 'B' type vouchers to people with special qualifications and skills; and 'C' type vouchers to other workers. Although this restricted immigration to some degree, it was counterbalanced by an increase in the women and children accompanying or joining the head of the family, particularly among the West Indians. However, a high proportion of Indians and Pakistanis are bringing their children, which appears to indicate that they plan to settle here (B.B.C., 1965, p. 24). A later Act (1965) restricts the number of vouchers to a total of 8,500 a year.

---

*A second similar article is being compiled of studies in the U.S.A. which appear to be relevant to the British situation.

The majority of immigrants are young and the birthrate has been estimated by the Economist Intelligence Unit (1962) to be 25 per 1,000 which is above the United Kingdom average of 18 per 1,000. This birthrate appears to be higher among the less educated within these groups. Waterhouse and Brabban (1964) in an Immigrant Fertility Study found that in their area there was a demand and need for a family planning service. The Five-Year Survey of Race Relations estimated at the end of 1964 that there were about 250,000 immigrant children in England, a considerable number born here and a large number immigrating after their parents. These children are still arriving.

The immigrants' immediate need on arrival is to find accommodation and this tends to occur in areas which already desperately require additional housing for its normal population. The schools in these areas are frequently short-staffed, overcrowded, and have inadequate buildings. Any population influx into these areas accentuates these difficulties. Because the Commonwealth immigrants are identifiable easily by colour they appear to be accused by some of the host population as the causes of these difficulties. It may be noted that in many areas the influx of Irish has been greater than that of coloured immigrants.

The social conflict, hostility and antagonisms engendered by the newcomers' arrival in Great Britain was best interpreted in terms of an 'Immigrant Host' relationship rather than as a racial situation during the 1950s. As Patterson (1963) describes it: 'In this approach colour falls into its rightful place as only one, albeit an important one, of many factors influencing and often complicating a particular immigrant host relationship'. The British situation is not strictly comparable with the American or South African situation where the colour bar has operated for generations in law, custom and individual attitudes.

In Britain, antipathy to outsiders, whether they be non-British or people from other areas of Britain, is probably a cultural norm, states Gorer (1955), although colour-conscious immigrants may interpret this antagonism to skin colour. While there is evidence of discriminatory behaviour towards immigrants, Patterson (1963), Little (1947) and Banton (1958) do not think this is based on widespread 'prejudice' as such, Little defining prejudice as 'an emotional, irrational, rigid disposition that is not amenable to reason or modification'.

Two major hypotheses are advanced to explain the motivations behind discriminatory behaviour prevalent in Britain in varying situations.

(1) *The colour-class hypothesis* set out by Little (1947) which was developed from the stable situation he found in the Cardiff dock area in 1946 when British society was thought, due partly to its colonial past, to identify coloured people with the lowest social class. These people feel that they will lose or jeopardize their social status by associating with a coloured person. Colour has the same socially inferior connotation as English spoken ungrammatically.

(2) *The stranger hypothesis* is supported by Banton (1958), in which he feels that the coloured immigrants are seen as the archetypal strangers, both in appearance and behaviour, in a society where mild xenophobia is a norm. He insists that British norms of conduct relating to coloured people differ from other strangers only in degree, although there are other factors which allocate the coloured people to the end position on the scale of strangeness.

Banton (1958) did a survey on British racial attitudes in 1956 in which 300 people were interviewed throughout the country. He found that colour prejudice did not appear to be widespread, but that evidence of discrimination was undeniable. None of the interviewees felt that there were any biological differences relevant to questions of social equality, although they showed themselves appreciative of cultural variations. In particular, Banton found that only 10% of those interviewed thought that coloured people were inferior to white and this was so often qualified that he believed that only 2% of the sample endorsed it wholeheartedly.

Richmond (1955), using the results of a still unpublished survey of British attitudes towards coloured 'colonials' conducted in 1951, asserted that the population could be divided into three groups of roughly equal size, according to their prejudice. These groups were: one-third extremely prejudiced, one-third mildly prejudiced, and one-third tolerant. After criticism by Banton (1959), Richmond (1958) modified his position, conceding that attitudes in Britain are probably more favourable and less rigid than he had supposed. The results of Banton's own smaller sample survey of attitudes in six British towns would suggest that only one in twenty-five of the British population would subscribe to the doctrines of racial superiority, although this would not mean that all of them were rigidly prejudiced by our definition nor that they were prepared to translate such opinions into action.

Glass and Pollins (1960) conducted a survey and concluded that most attitudes toward coloured people were severely prejudiced among only 10% of the sample and that about 15% were extremely tolerant. Between these two the majority expressed some degree of prejudice. Collins (1957) in his comparative study of such ethnic and religious groups as Negroes, Moslems and Chinese in three regions of Britain found that there are great regional variations in racial attitudes. The reasons for these differences of attitudes appear to be complex. Factors such as the type of immigrant group, the time span of the immigrant settlement, the size and pattern of the immigrant group and the historical background of the host community all seem to play their part, says Collins.

Other studies dealing with coloured immigration in England are by Griffith *et al.* (1960), a summary of information on housing, employment and an appraisal of sociological literature on the subject; Davison (1964) brings this up to date; and Desai (1963), an anthropologist, studies Indian immigrants, their settlement and integration patterns and their associations. The B.B.C.'s *Colour in Britain* (1965) is an enlargement of an adult education broadcast series, which contains much useful, well-presented factual material not found elsewhere about immigrants in Britain, their problems and what action can be taken to help them and ourselves, through community co-operation.

Huxley (1964) sees prejudice as a culture conflict in outlook and customs. Ruck (1959) discusses the problems and home background of the West Indians. Aurora (1960) describes Indian workers in Southall, and George (1960) the Cypriot community in London. Israel (1965) has done a detailed survey of the immigrant community in Slough which provides much interesting material for further detailed study.

There are many who now feel that the attitudes quoted by Banton, Patterson and Little were representative when there were relatively few non-white Commonwealth immigrants, but that they are no longer applicable. With the increasing flow of immigrants from Commonwealth countries and their concentration in urban areas reaction against them began to mount, as the Notting Hill and Nottingham disturbances appear to indicate (Jephcott, 1964), and action was demanded to restrict immigration by various pressure groups.

Deakin (1965) points out that a Gallup Poll taken in June, 1961 found that only 21% of the sample were in favour of free entry as opposed to 67% in favour of restrictions, and 6% for a total ban. The Commonwealth Immigration Act of 1962 was passed with Labour opposition. In 1964, however, the Labour Party conceded the need for restriction by approving its renewal the following year. In July, 1964, a new Gallup Poll showed only 10% of those questioned in favour of free entry, 68% in favour of some restriction and 20% for a total ban. Immigration and race appeared to be a major issue for the electorate in a number of constituencies, six of which, including Smethwick, are reported in Deakin's book.

It appears that public attitudes are becoming more vocally prejudiced against the 'coloured' Commonwealth ethnic groups. N'Dem (1953), a Nigerian, in his study of white and coloured relationships in Manchester, found that the factor of skin colour permeates the entire range of social relationships and that the coloured person feels that all whites are basically prejudiced against black people. Colour prejudice and discrimination is, claims N'Dem, a fact in Britain, and the sooner it is recognized, or accepted, the easier it will be to achieve social justice and fair play. This point of view is also advocated by Hill (1965) on the basis of an attitude survey carried out among 120 people representing a cross section of householders, employers, employees and landladies in three areas of London whose densities of coloured immigrants varied from high, medium, to low. Hill suggests that there has been a considerable increase in colour prejudice in Britain since Banton's estimates. Working on the basis of a social distance scale indicated by four degrees of relationship, Hill described 20% as severely prejudiced and 30% considerably prejudiced. The 1965 White Paper, *Immigration from the Commonwealth* (Home Office, Cmd. 2739) was considered by some in all the political parties to be a concession to increasing hostility to coloured immigrants. A principal point of criticism is that it institutes a quota of 8,500 Commonwealth immigrants per year holding work permits. The controls are thought by some to be severer than those imposed upon aliens; they restrict entry to children of immigrants between 16-18 years, except in cases of hardship, and give priority of entry to professional people, who are most needed in their own country.

The Commonwealth Immigrants Advisory Council was set up by the Act of 1962 to advise the Home Secretary on matters affecting immigrants and to examine the arrangements and powers of the local authorities to deal with the welfare of the immigrants. Its four reports [Home Office (1963-1965)] during the period of its existence were on Housing (Cmd. 2119, Cmd. 2796), Education of Commonwealth Immigrants (Cmd. 2766) and on Immigrant School Leavers (Cmd. 2458). This Council established a National Committee for Commonwealth Immigrants in April 1964, with a paid advisory officer to collect and circulate information to local authorities on how best to improve relations between immigrants and the community. Under the White Paper on Commonwealth Immigration (Home Office, 1965) a new Committee for Commonwealth Immigrants has been established to replace the previous Council and Committee. It is headed by the Archbishop of Canterbury and will have the same Advisory Officer, Miss Nadine Peppard.

To keep abreast of all activities concerned with the varied ethnic groups both in Britain and throughout the world readers should refer to the monthly *Newsletter* published by the Institute of Race Relations since 1960 and their quarterly journal, *Race*. A monthly summary of new items from national and local newspapers relating to immigrants in the United Kingdom is published by the Economist Intelligence Unit called *Commonwealth Immigration.*

As it is impossible to review all the relevant literature in the field, we include several bibliographies in our own bibliography. The most recent, by Grayshon and Houghton (1966), is the most complete and includes a selection of works from the U.S.A. which are relevant to the situation in Great Britain.

II. General survey of the educational situation of immigrant children

Since the last decade, immigrant children have been entering schools in the Greater London region and in Midland and Northern cities in increasing numbers. In 1956, it was estimated that 8,000 pupils in London schools were of overseas origin. By 1964, the figure had risen to 38,000. The rate of increase has made the task of integration more difficult. In Bradford, a city of high immigrant population, the 1,500 children in September 1964 had risen by almost one-third, only six months later. In areas of high immigration density, immigrant children can form large percentages in some schools. For example, in Birmingham, although the average proportion of immigrant children is only 5% of the children in all maintained schools, there are 22 primary schools with between 16%-70% coloured immigrant children. This may cause racial tension, since parents are afraid that their children will be affected by the apparent lower standards of the immigrant children as the latter's numbers increase.

The local authorities have tried to meet the administrative, educational and social problems of this sudden influx into schools, already overcrowded and understaffed, in a variety of ways. The Ministry of Education Pamphlet No. 43 (1963), discussed the problems schools would meet with these children, advising flexibility in cultural matters and outlining some administrative procedures in teaching the children. These recommendations were not and could not have been based upon research, since research evidence was not available.

Education authorities have described their methods of teaching immigrant children in articles in the *Times Educational Supplement* by Sampson (1963), Burgin (1964) and Robin (1965). Most schools do not separate immigrant from English children at the primary stage, although they may have special teaching in the secondary schools for one hour a day. This occurs, for example, in Nottingham, Smethwick and Slough. Although immigrant children are not being segregated they are generally in the 'C' streams or remedial classes, because of their linguistic disabilities. This may not be due to low ability, since few authorities have evaluated their ability levels. Tapper and Stoppes (1963) point out that about half the immigrants at one secondary school are placed in the 'D' stream in the school each year. However, there was a tendency for them to move upward as their English improved. Arithmetic was reported to be a favourite subject. Immigrant children were well represented among Maths and Science prizewinners. Bolton, as Sampson (1963) describes, attempted to evaluate, by non-linguistic tests which they were aware were not culture free, the ability levels of pupils whom they thought might be able to profit from a grammar school education. These revealed an I.Q. range of between 76 and 131. Other means of assessment have been used, including oral tests and extended interviews.

Education authorities who make special provision for children who do not speak English have classes organized in different ways. Specialist visiting teachers are employed in Birmingham. They visit fifteen specialist classes who meet in one school and pupils are drawn from three schools. Other smaller groups (about 60 in number) are visited three times a week within the schools. The children participate with English children in the main part of the school. Bradford has been distributing its non-English-speaking Asians to full-time language tuition centres. In Southall, the practice of dispersal of immigrant children to schools outside the children's locality by school bus has been in effect since 1963, whenever the local school has over a certain percentage of immigrant children. Of its 20 schools, 14 have reception classes where children remain from twelve to eighteen months. Each class has a Welfare Assistant. Manchester Education Authority has appointed part-time teachers in their junior schools. The children are withdrawn from their normal classes for short periods of intensive instruction in English and number work each day. Yet in Nottingham, with a large immigrant population, 27 out of 113 primary schools with more than 10% immigrant children cater for linguistic needs by using group methods within normal classes.

The Inner London Education Authority has three centres where a special staff member teaches non-English-speakers from secondary schools. One of them is Battersea where the children attend three times a week for half a day at a time, whose organization Lee (1965) describes in detail. Many other London secondary schools have their own special arrangements to meet the widely different needs of their immigrant pupils. The primary school immigrant children are catered for in their own schools with the help of off-quota teachers.

Huddersfield, until September 1964, channelled all of its immigrant school population (2·8%) into Spring Grove School, which had a special English department. Children were taught conversation, reading and writing in several stages and joined the rest of the school for subjects such as P.T. and needlework. When sufficiently well-established they joined the appropriate primary or secondary school near their homes. This reception centre technique has been adopted by other boroughs such as Bradford, Bolton and Slough. However, in September 1964, Huddersfield had first to establish another special English department at a local secondary school and more recently as the number of immigrant children have increased further, primary and secondary schools have established special English classes. Spring Grove is still a reception and distribution centre.

Burgin and Edson (1966) describe the work of this school, the methods and material used, the educational and social problems and their assessment of future needs in immigrant education. The advantages and disadvantages of these and other methods of teaching the immigrant child are discussed by Jordan (1965) and Lee (1965). Other new studies which describe in greater detail methods used by different schools are by Hawkes (1966), a teacher with experience of teaching immigrant children, and another by McCarthy, headmaster of a primary school with 55% immigrant children, and co-author Arthur (Arthur and McCarthy, unpublished).

McColl and Marney (1965) have done a general survey of the observations and opinions of 91 teachers in Slough schools, which had a significant percentage of immigrant children. McCrea (1964) has also studied some of the problems of immigrant children and compared the methods used by different authorities to deal with them.

The Schools Council has sponsored a project conducted by Derrick (1966) to establish the number of immigrant children born overseas or here who have not been brought up with English as their first language, and the number of West Indian children and of children in other immigrant groups; and to find out the arrangements made by the local authorities to teach them. This feasibility study examines the extent and nature of the problem and makes recommendations for positive action that may be taken centrally or regionally, especially in the field of the provision of materials and curricula studies. This is a useful first step since many of the larger authorities have kept records of numbers of ethnic minority groups in the schools and of their standard of English.

The study and evaluation of the best methods of teaching English to immigrant children who speak no English or a dialect English, and the training of teachers in these methods should preferably be two separate projects. However, the great demand for teachers of immigrant children and the help needed by those already involved has necessitated programmes combining both phases. The Universities of Leeds and Edinburgh and some colleges have put on courses in the teaching of English as a Foreign Language. Researches which are being done in this field are described by the Linguistics Panel (1959) of the British Council.

The London University Institute of Education has a one-year course leading to a Diploma in the Teaching of English as a Foreign Language. In September it began an experimental one-term course for fifteen primary school teachers of immigrant children under the direction of Anne Blatch in Professor Pattison's department. It attempts to help teachers who are not language specialists to teach by the direct method.

The use of overseas teachers is being recognized as a valuable source of additional teachers for immigrant children by L.E.A.s and Colleges of Education. However, many of these need training before they can take their place in English schools. Dalrymple (1966) describes Barking's experience with a ten-week course. The Margaret McMillan Training College in Bradford began a similar fifteen months experimental course in January, 1966 and the Department of Education and Science is to start others at Nottingham and Wolverhampton later in the year. There is a certain amount of controversy about this issue, too, since some speech experts feel that the primary school children will acquire the wrong speech rhythms from overseas teachers.

In March, 1965 a group of teachers formed an Association for the Teachers of English to Pupils from Overseas whose work is aimed at tackling current schools problems, which stem from the central one of language. Working parties for specific projects have been organized and conferences and seminars have been sponsored by the Associatian on specific problems, as well as engaging in other related activities, and issuing a bulletin to members.

Books especially written for teachers of English to overseas pupils in England are few. Derrick's (1966) forthcoming book deals with the principles and methods of language teaching. Braziers *et al.* (1965) have produced Infant Readers prepared in co-operation with the Birmingham Education Committee.

The British Council's (1963) English Teaching Information Centre acts as an information centre for many aspects of teaching English as a second or foreign language. It publishes an English Teaching Bibliography (1963) and supplements (1965), which include brief reviews of new studies and textbooks and have sections in audio-visual aids and language laboratories. It has also originated a quarterly, *English Language Teaching* for those engaged in teaching English as a foreign or second language; the emphasis is for those teaching abroad. The Ministry of Education's Pamphlet No. 43 (1963) includes a brief list of books on the teaching of English as a foreign language and which provide some background knowledge of the countries from which immigrants originate. Textbooks have been published by Longmans and Collier-Macmillan (see references under 'Textbooks . . .', 1965, 1963) for teaching English as a foreign language.

Recently the Committee on Research and Development in Modern Languages (Cooper, 1965/66), under the British Council, has set up a sub-committee to deal exclusively with the teaching of English to speakers of other languages. Its first task is a review of research activities concerned with the education of immigrant children and adults in Great Britain. Once this information is obtained, it plans to make recommendations about further work likely to help the teaching of English to speakers of other languages.

There are many other different problems which need recognition and research. The London Head Teachers' Association's Memorandum (1965) makes many recommendations, and emphasizes that: (1) cultural barriers to learning need to be studied; (2) methods of mental and intelligence testing for immigrant children need to be developed; (3) investigations into what constitutes disturbed behaviour for the different ethnic groups should be made. The Nottingham Director of Education, W. G. Jackson (1965) feels that there is evidence of a higher incidence of deafness and maladjustment among immigrant children, but little data is available.

Whether or not a quota of non-English-speaking children should be established in each school or class is a controversial issue, on which there is as yet no guidance from research in this country. Special teaching arrangements for language teaching, for example, should perhaps be geared to the proportions of non-English-speaking children in the area and some recommend that the children are not isolated for too long periods from English-speaking classmates.

Commonsense would support the recommendation of the Department of Education and Science that the proportion of immigrant children taught in any one school should not be unduly high (Circular 7/65). However, fixing any particular proportion will not of itself solve the problem. Research studies in the U.S.A. suggest that integration will depend much upon staff attitudes, the programmes of the school and attitudes in the wider community outside. The danger inherent in fixed proportions and in enforced dispersal is that it may be perceived as colour discrimination.

III. Specific problems concerning the education of coloured immigrant children

A. *Problems in Determining Ability Levels*

Ferron (1965) in his survey of the results of researches on the test performance of 'coloured' children describes the advantages and disadvantages of the different types of tests used. He finds that research indicates that where circumstances exist which ensure that white and coloured groups have a common way of life, a common language and equal opportunity, as among the oriental groups in the western coast of the U.S.A. the differences in I.Q. are small and non-existent. However, the existence of colour prejudice and other social pressures may affect intelligence test scores as a result of differences in emotional and motivational climate.

Where the differences in environmental opportunity between children in advanced civilizations and those in some underdeveloped areas are not too great, the differences in ability scores can be expected to be correspondingly smaller as the Porteus (1930) Hawaian and American study shows. Where cultural differences and linguistic handicaps and differences in environmental opportunity are great, intelligence test scores are lower among children in non-technological societies than among those brought up in a West European or North American culture, which is industrial.

Vernon (1961) studied the intellectual development of a sample of West Indian schoolboys in the British West Indies in 1960. He compared them with a representative sample of English schoolboys in a later study (1965), and found that the former were retarded by environmental handicaps. Although the West Indian scores indicated only a moderate degree of retardation, the deficit of West Indian boys varied considerably with different tests and was most noticeable in practical and non-verbal 'g' tests, and much less serious for educational attainment tests which seemed to be based upon drill and rote learning. His major finding was that the factors which handicapped the children's mental development were 'poor socio-economic, cultural, and linguistic environment, defective education and family instability'. The major factor apparently influencing children's performance in 'g' and verbal tests was the cultural level of the home, parental education and encouragement, reading facilities and probably speech background.

Thus, it seems evident that although non-linguistic tests are more suitable, children from non-technical cultures are seriously handicapped in picture and performance tests as well.

Although tests for non-Western peoples are available (Couch, 1963), it is highly improbable that culture-free tests or culture-fair tests can be devised, yet if such tests were possible they would be less useful for cross cultural comparisons. Vernon argues that since the new nations and their citizens are aiming to achieve Western-style technical civilizations it is important that they should produce more individuals with Western-type mental skills.

Vernon (1965b) outlines nine major environmental handicaps to mental development, which are of importance in understanding whether we can help non-Western children achieve their potential and whether we can measure their potential. A recent study on educational development among Canadian Indians and Eskimos (Vernon, 1966) concludes that the whole pattern of a culture determines the educational and vocational potential and combines to reduce the effective intelligence. Ferron (1966) infers that differences on test scores at pre-school level between two groups of nursery school children, one a class of Freetown Creole children and the other a class of English children, are due mainly to environmental differences. Ferron (1964) summarizing further studies of West African children, lists cultural factors which contribute to test performances and which could be controlled.

Most studies would seem to indicate that test quotients are to a large extent dependent upon the degree of assimilation of Western culture and that poor test performances of coloured children are the result of emotional and motivational problems experienced in trying to reconcile two cultures and traditions.

Alleyne (1962), comparing intelligence and attainment of children in London, Wales and Trinidad, whose mother tongue is not English, shows that a very high percentage of bi-linguals born outside Britain are undeniably handicapped in I.Q. tests, especially verbal ones, and that even non-verbal intelligence tests would not be really satisfactory, for comparison with English-speaking mono-linguals.

Saint (1963) attempted to assess the scholastic problems faced by 100 Punjabi boys in secondary schools. Seventy-six of these had been in England less than two years, and 49 less than one year. Of those who could attempt the tests, numbering 40, almost all lacked aural-oral skill in English and this contributed to their retardation in reading and to a lesser degree in arithmetic. Although the mean I.Q. of the group was 15 points below the normative sample, it cannot be regarded as evidence of inferior intellectual capacity. That cultural factors are important determinants of performance in these tests appears to be indicated by a moderate positive correlation between I.Q. and length of residence and schooling in the Western environment.

Houghton (1966) has published a very interesting study at infant level on relative performance on intelligence tests by two groups of matched children, West Indian and English. He found that there was little difference in mean scores: the Jamaican group was 90 and the local English group was 92. These results are similar to, although much lower than, those found in a similar American study by Goodman (1964). Houghton advances the hypothesis that these depressed British scores might best be explained in terms of deprivation—social, linguistic and environmental. Linguistic difficulties appear to be an important factor in the development of the West Indian child as Baker (1965) makes clear in his study of immigrant junior school children. Burgin and Pickup (1964) have also made a study in the problems of testing of abilities of non-English-speaking children at their primary school.

B. Linguistic Difficulties

It appears, therefore, that language is one major factor in this 'culturally induced backwardness' of immigrant children and affects assessment of ability and actual school performance. The educational centre of gravity of the problem is most likely in this area. Linguistic limitations may be present in varying degrees and be covered by the following categories:

(1) *Total language deficiency*, where not only is a foreign language spoken, but the written script is alien.

(2) *Partial language deficiency*, where some, but very little, English is spoken in the home, or the child has acquired some English from a longer residence here. The vernacular script may or may not be based upon the Western alphabet.

(3) *Dialect impediments.* Some children may speak English fluently, but dialect interposes, or a 'pidgin' English is spoken so that problems of listening, interpreting, and later reading and writing are present. This is a particular problem for West Indians, where Creole dialects are evident.

Alleyne (1962, 1965), in his study on the effect of bi-lingualism on educational attainment, compared bi-lingual groups with a control group of matched mono-linguals. He found that on three objective tests of educational achievement the mono-linguals were significantly superior to bi-linguals. Even when the experimental and control groups were matched on non-verbal ability tests, it was found that bi-linguals were inferior to mono-linguals on the verbal ability tests. There was clear evidence that bi-lingual deficiency was due to the experimental group's lack of involvement in English life and culture. It was meaningful that native bi-linguals were not found to be significantly inferior to their paired mono-linguals whereas the foreign bi-linguals showed a distinct inferiority. The greatest danger Alleyne discerned was that the child learns the language for a certain segment of the environment in which he will operate and another language for other specific situations.

Alleyne's findings emphasize the need to develop reading skill, to develop an oral approach and the necessity of learning the language in relation to real life situations. The field of dialect English, which is used by the majority of West Indians, is being studied on an intensive scale by the Institute of Education at the University of the West Indies, and is reported by Jones (1965). The needs and problems of teaching English as a foreign language (Standard English) to those speaking the West Indian dialect is also outlined by Jones (1966). Although some research has been done (Cassidy, 1961; Cassidy and Le Page, 1966) more studies have been mounted to ascertain the languages at present in use, as a basis for teaching a second language.

The Conference Commission on Communication (Jones, 1965) concludes: 'The dialect or Creole English is an immature language which is clearly inadequate for expressing the complexities of present-day life, for complete understanding of human motivation and behaviour'. The work of Bernstein (1960/61) applied to public and formal language has a great deal of significance for the understanding of our immigrant children's linguistic problems, in particular those of the West Indian ethnic group.

Studies of language problems of immigrant children—both Asian and West Indian—have been done by Cavage (1965) and Craig (1963). Baker (1965) has studied language problems in a junior school and Wallis (1963) describes methods and materials used. Bell (1966) is studying the English linguistic patterns of Indian immigrants in Smethwick, which may have relevance to children's speech patterns. A very thorough study of bi-lingualism in education which has good recommendations for future research, arose from an International Seminar on the problems in this field (Department of Education and Science, 1965).

C. *Cultural Factors affecting the Education of Immigrant Children*

How successfully the school accepts its secondary role as an agent of socialization for English children and acculturation for the immigrant child is dependent, as Williams (1966) points out in her study, upon:

(1) Aims and attitudes of the teachers;
(2) the affinity of the parental attitudes with those of the school;
(3) the attitudes and aspirations of the immigrant children regarding themselves and their relationships to children of the host country.

Williams' study (1966) showed that the teachers in her study saw their role as 'socializing', tidiness and respect for authority, and'anglicizing', that is conforming to middle-class English behaviour patterns and aspirations. 'Integration' is expected to occur automatically if multi-racial children are educated together. Unfortunately, integration in the sense of tolerance and acceptance by the host group of the differences shown by the immigrant group does not necessarily follow, as most of our English studies indicate.

Although the effect of attitudes and prejudices of the children, English and Immigrant, to one another upon educational achievement has not been studied as such, several studies on ethnic prejudice, as revealed by attitudes and choice of friends, have been undertaken. Hubbard (1965) made a study of 5-6 year-old children in a suburban middle-class school and in an urban working-class school and found that young children are aware of colour and disadvantages in a play situation. Possibly due to the design of the experiment, he could not find any correlation between children's and parents' attitudes.

A study of the factors influencing friendship choices among school children in a very mixed racial area of Liverpool was done by Silberman and Spice (1950), although this is now somewhat outdated. It found that social and economic class (as represented by standards of dress) were more important in determining friendship than colour. However, in this area most coloured children had English mothers and the coloured group contained more than one-quarter Chinese, who were the best clothed, which, as the authors state, introduces a qualification of this result.

In a more recent study, Kawwa (1963) found more prejudice in a comprehensive school in Islington, which contained a minority of coloured and Cypriot children, than in Lowestoft, a town which had few immigrant inhabitants. He found that members of the three ethnic groups chose more of their members than is expected by chance. Prejudice was much more intense against the Cypriots than towards the coloured immigrants, probably due to the fact that they were the largest ethnic group in the school and area. Saint (1963) found a strong social cleavage between the majority English and the minority Punjabi groups in the classes studied. The in-group preference of the majority group was markedly stronger than the in-group preference of the minority group. The almost complete lack of English boys' choice of any Punjabis may be an indication of prejudice.

Kawwa (1965) studied 750 children, mainly English, with almost equal numbers of such minority groups as Cypriot, West Indian and Asian, in one primary and one secondary school. Ethnic prejudice was found to exist in children as young as seven years of age, and was found to be greater towards the coloured children. The majority of 7-8 year-olds were aware of nationality, country and colour, although awareness of differences was confused until about 10 or 11 years. In secondary schools ethnic cleavage appears to be significant in choice of friendship patterns and rejection patterns, and awareness of colour, national origin and religion is universal. The great majority of English children scored between (1) score for the most prejudiced, to (3) score for neutral attitude. All non-English European children seemed to achieve assimilation rapidly. In contrast, coloured children had a difficult time and probably will have as long as colour is an important sign of alienism. Prejudice among children came out in Kawwa's study, in their stereotyped opinions about different ethnic groups: 'They take our jobs, our homes; they are dirty'. These seem to be echoes of parental attitudes. The attitudes of the immigrant children generally were that the English were not friendly and did not believe that all men were equal; and 95% of the families had no English friends.

In Williams' study, Indian and Pakistani parents' attitudes towards acculturation to the English culture were negative. The stereotypes given by the children in Kawwa's (1965) study seemed to echo English parents' attitudes.

The attitudes and aspirations of the immigrant children to themselves must be viewed against the background of the cultural factors which the immigrant child must adjust to before he can benefit educationally from the school; these have been outlined by Varley and Grayshon at a conference sponsored by the London Council of Social Service (1964). Varley points out the difficult psychological adjustments which have to be made by a child who rejoins parents, who are by now relative strangers in a totally strange land.

During the period when he is adapting to this experience, perhaps with resentment or shock, he is sent to school. The school in England, as Grayshon points out in the case of West Indian and Pakistani children, for example, is completely different from the schools they have attended, probably for shorter periods in relation to their age than children here. The teaching methods and techniques, the use of books, the discipline must all be adapted to. The schools are so busy coping with language problems that they appear to have insufficient time to help the child or teenager explicitly to come to terms with his new environment. The young child may adapt with less stress than the young teenager. Patterson (1963) reported that it is the younger or second immigrant generation which seems to have the major problem of adaptation and acceptance. Feeley (1965) has studied the social integration of coloured immigrant children in selected secondary schools, and Henderson (1961) discusses the treatment of prejudice in contemporary education.

The young immigrant may find it difficult to bridge the gap between the two cultures and resort to a behaviour pattern of social aloofness or rebelliousness. As Kaushal points out in the Birmingham U.N.A. Conference Report on 'Racial Tensions Among Youth' (1965), there is an urgent necessity for some agency to help him with these problems. Should he achieve some adaptation, he will have greater expectations and yet he will be frustrated by being identified through colour by the host society, with the older immigrant generation, and with all the stereotypes held about it.

As Williams' (1966) study also indicated, the immigrant children face conflict at home since the standards of their ethnic culture clash with those of the English culture and often with those of their peer group.

Frenkel-Brunswick (1947) points out that in the course of an attitude survey, studying prejudice in children from 10 to 17 years of age in the U.S.A., she found that prejudice in children is often linked with the child's attitude to authority. She suggested that forced submission to authority appears to produce surface conformity, but that this represses violent feelings, which will often be vented on a scapegoat or minority group.

Other studies, on attitudes and prejudice, which have not yet become available, are those of Pushkin (1966) which compares the development of prejudice in the young child from 3-7 years old from three areas in London; of Akram (1963) who studies the moral judgments of Indian and Pakistani children; and of Pawaroo (1966) who studies group relationships in some of the multi-racial schools in England and in British Guiana. Dosanjh (1966) is studying the educational and social difficulties of Punjabi children in Nottingham and Derby schools.

The opportunities available for skilled and clerical work, particularly for those immigrant school leavers who have the ability to do so, is relevant to school adjustment. In Tapper and Stoppes' (1963) study the latter claims that many of their able West Indian boys constantly meet with discrimination in job seeking, if not from employers then through the employees. The Commonwealth Immigrants Advisory Council 'Immigrant School Leavers Report' (1964) suggests that on the evidence now available there is some reason for concern about the job prospects for boys and girls with above average ability. Figueroa (1966) is studying the work experience of West Indian school leavers.

D. *Social Education in a Multi-Racial Society*

This requires projects to attempt to educate both English and Immigrant children and the host community for the society we envisage in the future, especially society in which there is knowledge, understanding of and mutual respect for other cultures.

Although there are a number of American books and articles on this subject, we list here the main British and Unesco studies dealing with the teaching of tolerance. Bibby (1959) covers the facts about race and the ways in which it can be taught. Leach (1964) describes useful techniques for use in school and has a comprehensive bibliography for those who wish to obtain more detailed information. James and Tennen (1953) describe how far and in what ways good personal contacts with members of another race can change people and their attitude as a whole. This study describes how pupils aged 12-14 in two schools classes were taught for two weeks by two African teachers. Their attitudes were ascertained before and after the teaching. It was found that the visit improved attitudes towards Negroes and other coloured foreigners more often than could be expected by chance.

Himmelstrand (1964), discussing methods of educating against racial discrimination, states that no single method can be used to eradicate racial prejudice since it may have different sources and mean different things to different people. Domnitz (1962) contains programmes in human relations for schools and youth clubs; Domnitz (1965) also reports on an international meeting at the Unesco Institute for Education which discussed educational techniques for combating prejudice and discrimination and for promoting better intergroup understanding.

Davis (1963) presents a review and bibliography of selected research on attitude change which is relevant. Williams (1961) discusses changes in pupils' attitudes towards West African Negroes, following the use of different teaching methods. Polack (1957) deals with the promotion of teaching of race questions in primary and secondary schools in Britain.

As Gardiner (1965) pointed out in his first Reith Lecture: 'Issues such as the security of minority settlers in the newly-independent states, the migration of workers into the highly industrialized countries and the ways automation might effect their future, all pose problems which could easily inflame racial feelings'.

IV. Discussion

It is evident from the demographic figures quoted in this article that we have already become a multi-racial society. This is not apparent in rural areas, nor in most urban areas of Wales and Scotland, but it is very evident in the large cities and industrial areas of England. Teachers and educational administrators in regions where immigrants have settled in large numbers have become increasingly aware of this over the last ten years, and have improvised measures for dealing with the problems.

This survey highlights the fact that the educational problems which have arisen and the measures taken to overcome them, cannot be isolated from larger social issues. We are faced with a multi-dimensional situation, not with a narrowly-based school problem. In a sense, the plight of coloured immigrant children is essentially the plight of the indigenous socially disadvantaged children in slum and problem areas, but with the added dimensions of language difficulties and colour prejudice to overcome. This cannot be separated from questions of housing, health, facilities for pre-school and after-school activities, youth employment, and many other factors.

Many of the surveys listed indicate that of the three ethnic groups now settling in Britain—Indian, Pakistani and West Indian—it is the West Indian group which appears to illustrate our difficulties most vividly. Linguistically, they seem to have an advantage over the other two groups, but it is their partial mastery of English, and their use of dialect which seems to be a major barrier to educational motivation and achievement. They also appear to be more susceptible to the climate, in terms of health, and pose difficulties for the host community because of the high incidence of broken families or children of varied parentage. Much work needs to be done on the varied family structure of immigrants, since it is evident that authoritarian, or other climates within the family, affect attitudes to education.

Our bibliography, which is not exhaustive, indicates that an increasing number of writings, studies and research have been produced over the last ten years in Britain, as educators, and sociologists have become increasingly aware of a new situation. The first point to note is that most of this should be classified as 'literature' of a descriptive kind, based upon little else than direct and even subjective judgment. Some may be classified as studies, of a general survey type, which have sought to quantify data about coloured immigrants, but many of these have not been very rigorous in terms of design, methodology or evaluation. Those which can be classified as research, in the rigorous sense of the word, are very few indeed. Yet there is an urgent need for objective information, and an assessment of the many factors involved in the present situation.

A great deal has been written on the linguistic aspect of educating immigrants. This is a valuable area to explore, and the provision of classroom materials based upon sound research seems to us to be essential. Yet the provision of such materials, and the training of teachers to use them and to teach English as a second language is only one aspect. Even here there is a danger that not enough basic research will be done involving, for example, an analysis of the linguistic errors of varied ethnic groups, and that classroom materials and methods may be based upon superficial and subjective assessments.

A very simple measure would be to provide a continuous pooling of information by local authorities who have numbers of immigrants resident in their area. This is being done to some extent through national committees and through the Schools Council, but much more is still required. Two examples will illustrate this need. Some local authorities keep very full records, both in departments dealing with the social services and in schools. But some local authorities resist keeping school records which indicate ethnic group or country of origin, on the grounds that this will foster racial prejudice. The availability of this basic information is vital if research is to be engaged upon and if school authorities are adequately to discharge their duties. It will not foster prejudice, but possibly provide the data for helping our schools become better instruments of education for tolerance.

A further example is the manner in which the Department of Education and Science policy of a one-third dispersal of coloured immigrant children was initiated without the full implications being perceived. If a one-third dispersal policy is rigidly enforced, and if on the other hand immigrants tend to congregate in certain areas, then more and more schools will have to transport some children to other areas. Thus, we found that schools located in predominantly 'coloured' areas were often not being used to capacity since there were not enough 'white' pupils living in some areas to make up the necessary two-thirds proportion. The logic of this situation is that these schools will die, gradually shrinking in size, with more and more immigrant children going by bus to distant schools. Alternative policies are: a disregard of the dispersal policy and many almost totally 'coloured' schools; the transporting of 'white' children into these areas, which is done in

New York, but only in the face of great opposition; a carefully planned re-housing policy so that sufficient municipal housing is available, by the building of new houses and flats, to balance the 'white' population which might otherwise move away. Certainly, any of these solutions bristle with problems, but some research into the effects of a dispersal policy in education and their alternatives is needed.

Another area requiring systematic investigation is that of the assessment of abilities of coloured immigrant children. The surveys and research so far conducted indicate the inadequacy of the current means of assessing the abilities of other cultural or ethnic groups, as Vernon (1965) rightly points out. We suggest that while there is a culturally induced backwardness among these groups, we have allowed ourselves to think only of problems and not of potential. It is a possible hypothesis that although retarded, in an English educational sense, these children are *above average* compared with their peers in their country of origin, simply because their parents possess the means and initiative to emigrate. What we urgently need are instruments and measures, which will provide our teachers with indications of the potential of immigrant children, a venture which would require intensive, expensive and long-term research.

The fields of cultural difference and both cultural and racial prejudice have, as yet, been relatively unexplored in relation to the schools. A dimension which would seem to be important is how the varied incoming groups perceive, and misperceive, the host community and how perceptions can improve or deteriorate. Since learning itself, the growth of self-concepts and levels of aspiration are bound up with group identification this would appear to be very relevant to the school and classroom situation.

Finally, in surveying the literature we see that some clarification of aims in education for the coloured immigrant needs to be achieved. We feel that 'integration'—meaning that the incoming group adapts itself to permanent membership of our society in certain areas such as education, employment, legal and political fields and is accepted by our society as an entity differing in religion, culture and family patterns—should be the minimum aim! These children need a sufficient mastery of English to make full use of our educational system and through social studies should be helped to understand our culture. It seems vital to us to explore how best we can help both the British and the immigrants to understand their cultures, religions, and ethnic identity as enriching contributions within the British culture.

Should they wish 'to assimilate', which implies complete acceptance of our values and patterns of living instead of those of their own ethnic group, this should be their choice but not forced upon them! However, those who wish to assimilate find that they cannot achieve the corollary of full assimilation which is complete acceptance by the host society because of their obvious colour identity.

Bibliography

Akram, A. (1964). 'The Moral Judgements of Indian and Pakistani Children'. Unpublished Ph.D. thesis, Birkbeck College, University of London.

Alleyne, M. H. (1962). 'The Teaching of Bi-lingual Children: Intelligence and Attainment of Children in London, Wales and Trinidad whose Mother Tongue is not English'. Unpublished M.A. dissertation, University of London.

Alleyne, M. H. (1965). 'Research on the Effects of Bi-lingualism on Education', p. 60; *see* Jones, J. ed. (1965) in this bibliography.

Arthur, J. *and* McCarthy, D. (1966). *Integration of Children of Immigrants into Schools.* To be published.

Association for Teachers of English for Pupils from Overseas (January, 1966). *Secretaries' Notes.* London: D. Cowan and N. Hawkes, 37 Fellows Road, London, N.W.3.

Aurora, G. S. (1960). 'Indian Workers in England'. Unpublished M.Sc. (Econ.) dissertation, University of London.

Baker, G. (1965). 'The Language Problem of Immigrant Children in Junior School'. Unpublished study, Diploma in Primary Education, University of Nottingham, Institute of Education.

Banton, M. (1958). *White and Coloured.* London: Jonathan Cape.

Bell, R. (1966). 'The Grammar of the English Spoken by Indian Immigrants in Smethwick'. Unpublished M.A. dissertation in final draft, University of Birmingham.

Bernstein, B. (1958). 'Some sociological determinants of perception'. *Brit. J. of Sociol.*, vol. 9, no. 2, June 1958, pp. 159-174.

Bernstein, B. (1960). 'Social Structure, Language and Learning'. *Educational Research*, vol. III, no. 1.

Bibby, C. (1959). *Race, Prejudice and Education.* London: Heinemann.

Bibby, C. (1964). 'The Roots of Prejudice'. Noah Baron Memorial Lecture. London: World Jewish Congress, British Section.

Birmingham, United Nations Association (1965). Conference Report on 'Racial Tensions Among Youth'. Birmingham: United Nations Association Branch.

Brazier, D., Cooke, E. *and* Jones, E. (1965). *Living Together.* Books 1-10. Oxford: Pergamon Press.

British Broadcasting Corporation (1965). Edited by Hooper, R. *Colour in Britain.* London: British Broadcasting Corporation.

Brocklebank-Fowler, C., Bland, C. *and* Farmer, T. (1965). 'Commonwealth Immigration'. London: Bow Group, Conservative Political Centre.

Burgin, T. (1964). 'English for immigrants: a Huddersfield experiment'. *Times Educational Supplement*, 10 January, 1964.

Burgin, T. *and* Edson, P. (1966). *Spring Grove: An Experiment in the Education of Immigrant Children.* London: The Institute of Race Relations: Oxford University Press.

Burgin, T. *and* Pickup, K. T. (1964). 'Problems of Testing the I.Q. of Non-English-speaking children'. Unpublished study of work at Spring Grove School, Huddersfield.

Cassidy, F. (1961). *Jamaica Talk.* New York: Macmillan.

Cassidy, F., *and* Le Page, R. B. (1966). *Dictionary of Jamaican English.* New York: Cambridge University Press.

Cavage, P. K. (1965). 'An Experiment in Sampling the Spoken Language of Immigrant Children'. Dissertation for Post-graduate Diploma in English as a Second Language, University of Leeds.

Central Office of Information (1965). *The Commonwealth.* Cmd. 167351. London: H.M.S.O.

Collins, S. (1957). *Coloured Minorities in Britain.* London: Lutterworth Press.

Cooper, A. R. (1965/66). 'Review of Current Research on Education of Immigrant Children and Adults Project'. Sub-committee of the Committee on Research and Development in Modern Languages. London: English Teaching Information Centre, British Council.

Couch, M. (1963). 'A select list on intelligence and other tests as used with non-western peoples'. *Education Libraries Bulletin*, no. 18.

Craig, D. (1963). 'A Comparative Study of the written English of some fourteen-year-old Jamaican and English Children'. Unpublished M.A. thesis, University of London, Institute of Education.

Dalrymple, A. H. (1966). 'Preparing immigrants for the master's desk'. *Times Educational Supplement*, 4 February, 1966.

'The Dark Million' (1965). *The Times*, 18 January, 1965. Article in a series.

Davis, E. E. (1963). 'Attitude Change: A Review and Bibliography of Selected Research'. *Reports and Papers in the Social Sciences.* Paris: Unesco Series.

Davison, R. B. (1964). *Commonwealth Immigrants.* London: Oxford University Press.

Deakin, N. (1965). *Colour and the British Electorate.* London: Pall Mall Press for the Institute of Race Relations.

Department of Education and Science Circular, 7/65 (14 June, 1965). 'The Education of Immigrants'. London: H.M.S.O.

Department of Education and Science (1965). 'Bi-lingualism in Education'. London: H.M.S.O.

Derrick, J., Research Fellow (1966a). *Immigrant Children: Teaching of English Report.* University of Leeds, Institute of Education. Unpublished report.

Derrick, J. (1966b). *Teaching English to Immigrants.* London: Longmans.

Desai, R. (1963). *Indian Immigrants in Britain.* London: Oxford University Press.

Domnitz, M. (1962). *Learning to Live with our Neighbours.* London: Woburn Press.

Domnitz, M. (1965). 'Educational techniques for combatting prejudice and discrimination and for promoting better intergroup Understanding'. *Report of International Meeting, 30 May,* 1964. Hamburg: Unesco Institute for Education.

Dosanjh, J. (1966). 'Problems of Immigrant Children from the Punjab in Schools in Nottingham and Derby'. M.A. dissertation, University of Nottingham.

Economist Intelligence Unit (1962/63). 'Studies on Immigration from the Commonwealth', Pamphlets 1-5. London: Economist Intelligence Unit Ltd.

Economist Intelligence Unit (1964+). 'Commonwealth Immigration: A Monthly Summary'. London: Economist Intelligence Unit Ltd.

Edson P. (1966). 'East and West meetings', *in:* 'Education for Integration'. *Institute of Race Relations Newsletter.* London: Institute of Race Relations.

'Education for Integration', monthly series; Institute of Race Relations *Newsletter.* London: Institute of Race Relations.

Elton, Lord (1965). *The Unarmed Invasion.* London: Geoffrey Bles.

English-Teaching Information Centre (1963). 'English Teaching Bibliography'. London: British Council.

English Teaching Information Centre (1965). 'English Teaching Bibliography Supplement'. Parts I and II. London: British Council.

Feeley, M. R. (1965). 'An Investigation of the Social Integration of Coloured Immigrant Children in Selected Secondary Schools'. Dissertation for Diploma in Secondary Education, University of Liverpool, Institute of Education.

Ferron, O. (1964). 'A Study of Certain Factors Related to the Tested Intelligence of Groups of some West African Children'. Unpublished Ph.D. thesis, University of Edinburgh.

Ferron, O. (1965). 'The test performance of "coloured" children'. *Educational Research,* vol. VIII, no. 1, November 1965, pp. 42-57.

Ferron, O. (1966). 'The effects of early environmental stimulations among Freetown Creoles—a comparative study'. *Educational Research,* vol. VIII, no. 3, June 1966.

Figueroa, P. (1966). 'Work Experience of West Indian School Leavers'. Working title of a study in progress for Ph.D. degree at the London School of Economics, University of London.

Foot, P. (1965). *Immigration and Race in British Politics.* Harmondsworth: Penguin.

Frenkel-Brunswick, E. (1947). 'A study of prejudice in children'. *Human Relations,* vol. 1, no. 3, pp. 295-306.

Gardiner, R. (1965). 'First Reith Lecture'. *Guardian,* 8 November, 1965.

General Register Office (1965). 'Census 1961, England and Wales: Commonwealth Immigrants in the Conurbations'. London: H.M.S.O.

George, V. (1960). 'The Assimilation of the Cypriot Community in London'. Unpublished M.A. dissertation, University of Nottingham.

Glass, R. *and* Pollins, H. (1960). *The Newcomers: The West Indian in London.* London: Allen & Unwin.

Gorer, G. (1955). *Exploring English Character.* London: Cresset Press.

Green, C. (1931). 'Racial prejudice among school children'. *Geography,* vol. XVI.

Griffith, J., Henderson, J., Usborne, M., Wood, D. *and* Long, H. (1960). *Coloured Immigrants in Britain.* London: Oxford University Press.

Hawkes, N. (1966). *Immigrant Children in British Schools.* Institute of Race Relations. London: Pall Mall Press.

Henderson, J. (1961). 'The treatment of prejudice in contemporary education'. *Race,* vol. III, no. 1.

Hill, C. (1965). *How Colour Prejudiced is Britain?* London: Gollancz.

Himmelstrand, U. (1964). 'Content and Methods of Education in Dealing with Racial Discrimination'. London: The International Federation of Workers' Associations.

Home Office (1965). 'Immigration from the Commonwealth' (Cmd. 2739). London: H.M.S.O.

Home Office. Commonwealth Immigrants Advisory Council (1963-1965).
First Report (1963). *Housing.* Cmd. 2119. London: H.M.S.O.
Second Report (1964). *Education of Commonwealth Immigrants in British Schools.* Cmd. 2266. London: H.M.S.O.
Third Report (1964). *Immigrant School Leavers.* Cmd. 2796. London: H.M.S.O.
Fourth Report (1965). *Housing.* Cmd. 2796. London: H.M.S.O.

Houghton, V. P. (1966). 'A report on the scores of West Indian immigrant children and English children on an individually administered test'. *Race,* vol. 8, no. 1. London: Institute of Race Relations.

Hubbard, J. (1965). 'Racial Attitudes in Young Children and their Parents' Attitudes to Child Rearing and Social Problems'. Unpublished Child Development Report, University of London, Institute of Education.

Huxley, E. (1964). *Back Street: New Worlds.* London: Chatto & Windus.

Independent Labour Party (1965). 'Second Class Citizens'. London: Independent Labour Party.

Institute of Race Relations (1960+). Editor, Patterson, S. Monthly *Newsletter.* London: Institute of Race Relations.

Institute of Race Relations. *Race Quarterly.* London: Institute of Race Relations.

Israel, W. H. (1964). *Colour and Community: A Study of Coloured Immigrants and Race Relations in an Industrial Town.* Slough: Slough Council of Social Services.

Jackson, W. G. (1965). 'The newcomers: education of the immigrant child'. *Lancet,* 23 October, 1965.

James, H. E. O. *and* Tenen, C. (1953). *The Teacher was Black.* London: Heinemann.

Jephcott, P. (1964). *A Troubled Area: Notes on Notting Hill.* London: Faber.

Jones, J., ed. (1965). 'Linguistics, and Language Teaching in a Multilingual Society'. Report of Conference, 6th—9th April, 1964. University of West Indies, Kingston, Jamaica. Old Woking, Surrey: Unwin.

Jones, J. (1966). 'English in the Commonwealth: the West Indies'. *English Language Teaching*, vol. XX, no. 2. London: Oxford University Press.

Jordan, R. R. (1965). 'Integration of Coloured Commonwealth Immigrants within the English Educational System'. Dissertation for post-graduate Certificate of Education—English as a foreign language. University of London, Institute of Education.

Kawwa, T. (1963). 'The Ethnic Prejudice and Choice of Friends among English and Non-English Adolescents'. Unpublished M.A. dissertation, University of London, Institute of Education.

Kawwa, T. (1965). 'A Study of the Interaction between Native and Immigrant Children in English Schools with special reference to Ethnic Prejudice'. Unpublished Ph.D. thesis, University of London, Institute of Education.

Leach, P. (1964). 'Teaching tolerance: the role of the school in furthering constructive inter-group relations'. *International Review of Education*, vol. X, no. 2.

Lee, R. (1965). 'The education of immigrant children in England'. *Race*, vol. VII, no. 2.

Le Page, R. B. (1955). 'The language problem in the British Caribbean'. *Caribbean Quarterly*, vol. 4, no. 1.

Linguistics Panel (1959). 'University Training and Research in the Teaching of English as a Second/Foreign Language'. London: The British Council.

Little, Kenneth (1947). *Negroes in Britain*. London: Kegan Paul.

London Council of Social Service (1964). 'The Young Immigrant at Home and at School'. Report of a Day Conference, 19th May.

London Head Teachers' Association (1965). 'Memorandum on Immigrant Children in London Schools'. London: Pegg & Sons Ltd.

McColl, T. *and* Marney, B. R. (1965). 'An Investigation into some of the Factors affecting the Integration of Immigrant Children and Adolescents in a Selection of Schools in Slough'. Joint dissertation under Regulation 9(c), Borough Road College, Isleworth.

McCrea, W. (1964). 'A Study of some Problems presented by Immigrant Children in the Schools of one Authority with additional reference to the work being done by other Authorities'. Dissertation for Diploma in Education, University of Liverpool.

Ministry of Education (1963). *English for Immigrants*. (Pamphlet No. 43.) London: H.M.S.O.

Morrison, J. C. (1958). *Puerto Rican Study*, 1953-1957. New York City: Board of Education.

N'Dem, E. (1953). 'Negro Immigrants in Manchester'. Unpublished M.A. dissertation in Anthropology, University College, University of London.

Nottingham University, Institute of Education (1963). 'The Education of Immigrant Children'. Report of a Conference, *Bulletin*, no. 42, pp. 1-8.

Patterson, S. (1963a). *Dark Strangers*. London: Tavistock Publications.

Patterson, S., ed. (1963b). 'Immigrants in London'. Report of a Study Group (set up by the London Council of Social Service). London: N.C.S.S.

Pawaroo, M. 'Group Relationships in some of the Multi-Racial Schools in England and in British Guiana'. Unpublished Ph.D. thesis in final draft, University of London.

Peach, C. (1965). 'Immigrants and the 1961 census'. *Newsletter*. Institute of Race Relations, 10 October, 1965.

Polack, A. (1957). 'The promotion of teaching of race questions in primary and secondary schools'. *Anglo-Jewish Association Journal*, vol. 3, no. 2.

Porteus, S. (1930). 'Race and social differences in performance tests'. *General Psychology Monographs*, 8, no. 2.

Pressure for Economic and Social Toryism (1965). 'Immigration for the Commonwealth'. London: Conservative Political Centre.

Pushkin, I. (1966). 'An Investigation into the Development of Prejudice in Young Children'. Unpublished Ph.D. thesis, Birkbeck College, University of London.

Richmond, A. (1955). *The Colour Problem*. Harmondsworth: Penguin.

Richmond, A. (1958). 'Recent research on race relations in Britain'. *International Social Science Bulletin*, vol. X, pp. 344-72.

Robin, C. E. (1965). 'Handling the immigrants: English classes at Smethwick'. *Times Educational Supplement*, 1 January, 1965.

Ruck, S. K., ed. (1959). *The West Indian comes to England*. London: Routledge.

Saint, C. K. (1963). 'Scholastic and Sociological Adjustment Problems of the Punjabi-speaking Children in Smethwick'. Unpublished M.Ed. dissertation, University of Birmingham.

Sampson, O. (1963). 'Eleven-plus for immigrants'. *Times Educational Supplement*, 5 April, 1963, no. 2498.

Shaw, A. B. (1966). 'Rural education in the Punjab', *in:* 'Education for Integration', Institute of Race Relations *Newsletter*. London: Institute of Race Relations.

Silberman, L. *and* Spice, B. (1950). *Colour and Class in Six Liverpool Schools*. Liverpool: University Press.

Tapper, O. *and* Stoppes, E. (1963). 'Educating the immigrant'. *East London Papers*, vol. 6, no. 2.

Textbooks for Teaching English as a Second Language:
  (1965). *Structural Readers*. Stages 1-6. London: Longmans.
  English Language Service (1963). *English this Way*. Books 1-12. London: Collier Macmillan.

Vernon, P. E. (1961). 'Intellectual Development in Non-Technological Societies'. Proceedings of the 14th International Congress of Applied Psychology, vol. II, 94-105. Copenhagen: Munksgaard.

Vernon, P. E. (1965a). 'Environmental handicaps and intellectual development'. *Brit. J. of Educ. Psychol.*, XXX, Parts 1 and 2.

VERNON, P. E. (1965b). 'Ability factors and environmental influences'. *American Psychologist*, vol. 20, no. 9.

VERNON, P. E. (1966). 'Educational and intellectual development among Canadian Indians and Eskimos'. *Educ. Rev.*, vol. 1, no. 1.

WALLIS, B. (1963). 'English for Immigrant Children in Batley'. Dissertation for post-graduate Diploma in English as a Second Language, University of Leeds.

WATERHOUSE, J. *and* BRABBAN, D. (1964). 'Inquiry into fertility of immigrants'. *Eugenics Review*, vol. 56, no. 1.

WILLIAMS, H. (1961). 'Changes in pupils' attitudes towards West African Negroes following the use of different teaching methods'. *Brit. J. of Educ. Psychol.*, vol. 31.

WILLIAMS, J. (1966). 'Immigrant Children: Schools' Role in Integration'. Working title for one study, which is part of a larger study by Rex, J. (1966). *Race, Community and Conflict: A Study on Sparkbrook.* London: O.U.P.

YOUNG FABIAN GROUP (1965). *Strangers Within.* London: Fabian Society.

YOUNG LIBERALS, National League of (1965). 'No Colour Bar Here'. London: National League of Young Liberals and Union of Liberal Students.

BIBLIOGRAPHIES on Problems of Immigration and Race Relations, mainly in Great Britain:

BRISTOW, T. *and* EARL, M. (1965). 'A Selected Bibliography on Immigration and Race Problems'. University of London, Institute of Education.

ENGLISH-TEACHING INFORMATION SERVICE (1965). 'Immigrant Children: Language, Education and Social Integration Problems'. London: Education Division, British Council.

GRAYSHON, M. C. *and* HOUGHTON, J. P. (1966). 'Initial Bibliography of Immigration and Race'. Occasional Paper no. 6. University of Nottingham, Institute of Education.

PATTERSON, S. (1965). 'References Bibliography'. *Dark Strangers.* London: Tavistock; and Harmondsworth: Penguin.

SIVANANDAN, A. (July, 1965). 'Coloured Immigrants in Britain: A Select Bibliography'. London: The Institute of Race Relations.

SIVANANDAN, A. (July, 1965). 'Race Relations and Education: A Reading List'. London: The Institute of Race Relations.

READING AND ADJUSTMENT: A REVIEW OF THE LITERATURE

by Olive C. Sampson

Lecturer in Education (Child Guidance), University of Manchester

I. Introduction

MANY class teachers and remedial workers must have made their private observations on the relationship between personality and adjustment, probably reaching different conclusions in the process. They may or may not agree that a significant proportion of retardates (that is, children functioning below their potential level), show an element of emotional disturbance. Their opinions are also likely to differ even more as to whether any maladjustment noticed is the cause or the effect of the scholastic failure. The present article seeks to review the literature relevant to these problems. First, I shall report on the evidence regarding the existence of some sort of association between reading achievement and adjustment. Next I shall consider researches and studies which throw light on the direction and type of the relationship. I shall also refer to the findings concerning the kinds of maladjustment that are common in connection with reading failure, as these help to indicate the psychological dynamics of the situation. Finally the whole tendency of the evidence will be discussed and reviewed.

II. Evidence concerning an association between reading and adjustment

Evidence of an association between reading achievement and personality adjustment comes from a number of sources. Among these the surveys are fundamental.

Interest in the subject appears to date from the first quarter of this century, with an article by Burt (1917) entitled, ' The Unstable Child '. A similarly seminal article in America was Grace E. Bird's (1927), ' Personality Factors in Learning '. Subsequently the whole question received much detailed attention in the States and important research was done, but the major contribution towards establishing the fact of a relationship is British. It comes in *The Backward Child*, first published in 1937. In this book Burt was exploring the causes of educational backwardness as revealed in large-scale surveys in London and Birmingham. These showed that most of such backwardness was linked with inferior intelligence. Other causal factors such as physical ill-health, poor home circumstances, irregular attendance, and various ' specialized disabilities ' explained most of the rest. But in 9% of all cases Burt found himself forced to consider the influence of ' temperament and emotion '. To investigate the situation further, the whole backward group was compared with a control group otherwise similar but successful in school work. Certain personality variables were assessed using individual psychological interviews and teacher's reports, with the result that a low but positive correlation appeared between backwardness and certain personality characteristics. The figures are given in Table 1.

Table 1

Percentage of Temperamental and Neurotic Conditions in Backward and Control Groups

	Control	*Backward*	*Correlation with backwardness*
Temperamental Conditions			
Unstable			
(a) excitable	9·7	19·9	.24
(b) repressed	5·1	9.0	.16
Lethargic	1.8	4·1	.20
Nervous Conditions			
Anxiety states	3·3	6.4	.17
Neurasthenia	·5	1.8	.26
Hysteria	·3	0.8	.19*
Minor symptoms	14.8	18.2	.07*

*Not statistically significant.

Various subsequent surveys, mostly in Midland towns between 1949 and 1956 (Collins, 1961) drew further attention to the phenomenon of retardation, and the idea that underfunctioning might be associated with maladjustment gained currency. In 1958, Stott quoted figures from surveys covering about 1000 children. Assessments of the children's adjustment had been made on the now well-known *Bristol Guides*. The results indicated that the more backward were also the more maladjusted. Morris (1959) investigating reading backwardness among Primary School children in Kent, asked 60 Head-teachers to estimate the causes of this. The analysis showed 42 ' centred on the school ', 62 ' centred on the home ', and 114 ' centred on the child '. Of these last, 23% were concerned with adverse adjustment. Further, in an article (1963) and in a paper delivered (1965) at the Annual Conference of the British Psychological Society, Chazan referred to a number of recent research studies showing a relationship between adjustment (again estimated on the *Bristol Guides*) and attainment. A relevant item from his tables is quoted below. It shows the mean *Bristol Guide* scores of third year Primary School children who are high and low achievers in reading. The difference is significant at 1% level.

Table 2

Primary Schools

	Mean Bristol Guide Maladjustment Score
Highest Readers	6.9
Lowest Readers	17.2

To sum up then, it is evident that surveys of various sorts reveal a tendency for reading difficulty and maladjustment to be associated. Other researches, large and small, examine this association more closely and to these we shall now turn.

III. Does reading failure cause maladjustment?

It is often accepted, as if it were self-evident, that reading failure, by lessening the child's confidence and increasing his anxiety, will be likely to cause maladjustment. The maladjustment may show in negativism, apathy, or in other ways. Monroe (1932) and Fernald (1943), both very experienced in remedial reading work, incline to this view. Fernald points to the evidence of a series of 78 children presenting serious reading problems. Of these, only 4 had shown any evidence of personality difficulties before entering school. Monroe, reviewing the situation, believes that failure in reading is ' more frequently ' the cause than the result of emotional and personality problems. She instances ' the tendency for remedial instruction in reading to improve not only the educational achievement but also to affect the related personality problems '. Preston (1940) likewise, quotes the evidence of results, citing failures ' who were badly maladjusted ' but who ' recovered on being taught to read '. She feels strongly about the situation. As the result of a study of 100 children in grades 2–10 with I.Q.s 90–140 who were poor readers, compared with a control group of adequate readers, she said: ' reading failure causes not only a blighting insecurity in the school world which gives rise to serious maladjustments—but also an embarrassing and belittling insecurity in the home '. Stewart (1950) and others would endorse her views on the importance of insecurity. However, this is not in his view inevitably associated with reading difficulty. According to his study of 30 maladjusted children insecurity was characteristic of them all, whether they were good or bad readers.

IV. Does maladjustment cause reading failure?

Less familiar to the average teacher is the idea that something in the child's personality, whether a temporary reaction or a permanent trait, may affect his progress in reading. There are, however, a number of studies into this possibility. One of the most outstanding is Edith Gann's (1945), *Reading Difficulty and Personality Organization.* In this she starts from the hypothesis that 'dynamic processes in the personality organization which determine its means and type of adaptation are related to and influential in the reading experience'. She examined three equated groups of children who were respectively retarded, average or advanced in reading. To measure their personality make-up she used the Rorschach technique, a clinical instrument in which she was expert. A variety of findings emerged but her general conclusion is that most relevant here: 'Retarded reading', she affirmed, 'should be considered a personality problem as well as a learning problem'. She considered that in certain cases, 'definite personality therapy' would be essential for improvement. This view is endorsed by some (Axline, 1947, Bills, 1950) who have actually used psychotherapy, without any special remedial teaching, with retarded groups. Others who have done case-studies (Blanchard, 1935—4 cases; Liss, 1935—1 case; Sylvester and Kunst, 1943—3 cases) and follow-up work (Castner, 1935—13 cases) also provide evidence of psychogenic factors in the reading situation. Comparable but on a larger scale is the work of Ellis (1949) and Young (1938). Young's children numbered 41, all patients of a Psycho-Educational Clinic. They were followed up from 3½ to 12 years after first referral. Young concludes: 'Perhaps the most significant factor resulting from this study is the evidence that the emotional reactions of parents and children are of the utmost importance' for learning to read. Ellis investigated 100 children referred to a Mental Health Clinic. They were all screened for reading disabilities and received remedial teaching, but more emphasis was put on the simultaneous Child Guidance approach. There were numerous home visits, therapeutic interviews and so on. Ellis concludes that 'reading disability, though specific enough in one sense, is really an organismic problem, and effective treatment would seem to involve the total personality of the child rather than some set of particular sensory or intellectual faculties'.

Chester C. Bennett's earlier study (1938) which is one of the most careful of its kind, introduces a different point of view from the above. 'Among the multiple possibilities there is no single gestalt of eteological factors which is characteristic of poor readers and not present in the experience of most normal readers'. His inquiry was however directed to a great variety of factors and it is worth noting that, in spite of the above remark, he found the poor readers had a significantly more adverse adjustment than the good towards the task of reading.

V. What sorts of maladjustment are involved in reading failure?

Referring to the possible causal relation existing between emotional adjustment and reading failure, Gates (1941) said: 'on scarcely no important professional issue in reading is there more disagreement'. This would also appear to apply when an attempt is made to identify the sorts of maladjustment most commonly involved. Many considerable investigations have been made in this area but the results are confusing.

Some give very general descriptions of the personality types revealed by their researches as characteristic of poor readers. Thus 'emotional instability' is stressed by Castner and by Gann. 'Highly-strung' personalities, evincing restless behaviour, are reported by Jackson (1944). Knoblock (1965) finds 'less mature' attitudes in the Rorschach records of poor readers. Others stress particular personality patterns, but identify absolutely contrasted types from the mainly aggressive and defensive (Spache, 1947) to the nervous and discouraged (Zolkos, 1951). Some, like Gates (1941), find ' no single personality pattern ' but admit all and identify ' types of situations producing personality maladjustments ' of various sorts. Challman (1939), who knows retarded children in the classroom, finds five types, which overlap with Gates', involving nervousness, withdrawal, aggression, defeatism and worry. A few (Ladd, 1933; Lantz and Liebes, 1943; Monroe 1932) emphasize maladjusted school attitudes rather than general maladjustment as important.

In this welter of opinions the most firmly based are those of Gann, Gates, Jackson, Ladd, Natchez, and Spache, in the sense that their work is sufficiently rigorous; but there is not much agreement between them. Gates' views have already been indicated, and Gann's research has been described for its other findings. In assessing the personality make-up of her retarded readers she finds them less emotionally stable, more insecure and fearful, and socially less adaptable than are the average and good achievers. Jackson compared 300 retarded and 300 advanced readers, using ratings; ' sub-average ' personalities were significantly more frequent among the former. Ladd (1933) in an excellent study, finds lack of confidence and lack of persistence characteristic of her 12 most retarded in comparison with a matched group of the 12 best readers from among 315 children reviewed. Natchez (1959) in an equally careful study of oral reading, associates frustration—expressing itself in excessive dependence, aggression or withdrawal—with poorer reading. Spache (1957) used the Rosenzweig Picture-Frustration Study with 100 retarded readers and found that, in comparison with the norms, they ' were inclined to be more aggressive and defensive than children of their age, less insightful, and are relatively poor in knowing how to handle situations in conflict with adults '.

VI. Discussion

The problem that now faces us is: Does existing research on reading and adjustment justify any general conclusions? Have we any incontrovertible facts? If not, are any tendencies discernible? Or is any more research wanted ?

These questions (except the last) are hard to answer. The research literature is not only very varied in scope, but it is also beset by semantic uncertainties. The most obvious of these occur in reference to the dimension of adjustment. Are terms such as ' emotional disturbance ', ' maladjustment ', ' personality defect ', ' unfavourable emotional attitudes ' and so on, to be equated? Even reading achievement is defined by different standards in different researches. Some use reading ages, others grade averages, others ' failure to improve ', as their starting point in ways which make comparison very uncertain. The varying chronological ages involved also present a problem. Retardation in reading has surely a very different significance in readers of different ages, yet the literature deals with cases from 6 to 16 years of age without recognition of this. Lack of any agreement in defining these three dimensions—adjustment, reading attainment and age—makes it impossible to parallel results.

Again, the instruments of measurement raise problems by their variety. Adjustment is assessed in very many ways—by teacher's ratings, by questionnaires filled up by children, by clinical techniques, by the accident of child guidance referral, etc. The imprint of the method chosen is again and again seen in the results. Reading tests in equal diversity are employed by different researchers. The categories of 'advanced' and 'retarded' readers sometimes denote the best and worst of a larger population—whether city, district, or school. At other times they denote those who are 1 or 2 years above or below grade expectation. When mental age and achievement age are the basis of comparison, the former is established in various ways, sometimes by the use of group tests, sometimes by the 'Draw-a-Man' test, at the other times by the individually administered Stanford-Binet scale.

As a result of all these methodological variations, detailed agreement between the findings of the different researches is not to be expected, and is certainly not to be found. But they have a common tendency. They agree, as Helen M. Robinson (1946) puts it, 'without exception, that a large number of severely retarded readers' also evidence 'emotional and personality maladjustment'. Different workers interpret the causal relation operating between adjustment and reading in different ways. Gates (1941), doyen of the field, puts the situation plausibly as follows: 'My estimate is that among cases of very marked specific reading disability about 75% will show personality maladjustment. Of these, the personality maladjustment is the cause in a quarter of the cases and an accompaniment or result in three-quarters'. Such a view does not rule out the concept of interaction. 'In many cases', says Robinson, who espouses this interpretation, 'educational maladjustments and reading failure would seem to interact, each adding to the seriousness of the other'.

In spite of the lack of conclusive evidence as to the dynamics of the situation, many writers when summarizing, seem to feel, with M. D. Vernon (1958), that there is a case for 'a thorough investigation of possible temperamental or neurotic defect' in every case of severe reading disability. For some this implies a need for psychotherapy: 'symptomatic treatment by pedagogical methods is not enough', state Sylvester and Kunst (1943), and a similar viewpoint is taken by Young (1938) when he remarks that 'without the appreciation of the dynamics of the emotions, tutoring . . . . may fall flat'. Common sense—and Gates (op. cit.)—come to the rescue with the latter's firmly stated view that at least 90% of retardation can be adequately dealt with by teachers, though 'skilful clinical investigation' may help the remaining ten in every hundred.

VII. What research is needed?

The reader of this survey may well conclude that while much of the work on reading and adjustment is suggestive, none of it is conclusive. The question therefore arises as to what further research is most needed. Desirable lines of inquiry would seem to be of three kinds.

(1) First the question of the amount of reading backwardness attributable to maladjustment requires careful investigation at different ages and at different I.Q. levels. Studies in different parts of the country would show any local variations in incidence if the same clearly defined terms and objective measures were used on sufficiently large and representative samples.

(2) Secondly, the nature of the maladjustments mainly involved in reading backwardness need study. The great variety of existing views on this question should be put to the test. Careful clinical investigations even with small but well defined groups

could tell us much. Objective assessments with larger populations could also usefully explore tendencies and perhaps reduce to acceptable dimensions the welter of opinion.

(3) Thirdly, the experimental approach of Axline and Bills could be taken further and applied with greater rigour to give a clearer picture of the emotional dynamics that may affect the reading situation.

Bibliography

**indicates most important research and review sources.*

AXLINE, V. (1947). 'Non-directive therapy for poor readers'. *J. Consult Psychol.*, 11, 61-69.
BENNETT, C. C. (1938). *An Inquiry into the Genesis of Poor Reading.* New York: Teachers Coll., Columbia Univ.
BILLS, R. E. (1950). 'Non-directive play therapy with retarded readers'. *J. Consult. Psychol.*, 14, 140-149.
BILLS, R. E. (1950). 'Play therapy with well-adjusted retarded readers'. *J. Consult. Psychol.*, 14, 246-249.
BIRD, G. E. (1927). 'Personality factors in learning'. *Personnel J.*, 61, 56-59.
BLANCHARD, P. (1935). 'Psychogenic factors in some cases of reading disability'. *Amer. J. Orthopsychiat.*, 5, 361-371.
BURT, C. (1917). 'The unstable child.' *Child Study*, 10, 61-79.
*BURT, C. (1946). *The Backward Child.* (pp. 538-563). University of London Press.
CHALLMAN, R. C. (1939). 'Personality maladjustment and remedial reading'. *J. Exceptional children*, 6, 7-11.
CHAZAN, M. (1963). 'Maladjustment, attainment and sociometric status'. *U. Coll. Swansea, Ed. J.*, 4-7.
CHAZAN, M. (1965). 'Maladjustment, Attainment and Sociometric Status'. Unpublished paper, Annual Conference B.P.S.
*COLLINS, J. E. (1961). *The Effects of Remedial Education.* Birmingham: University Institute of Education.
ELLIS, A. (1949). 'Results of a mental hygiene approach to reading disability problems'. *J. Consult. Psychol.*, 13, 56-61.
FABIAN, A. (1955). 'Reading disability—and index of pathology'. *Amer. J. Orthopsychiat.*, 25, 319-329.
FENDRICK, P. and BOND, G. (1936). 'Delinquency and reading'. *J. Genet. Psychol.*, 48, 236-243.
*FERNALD, G. M. (1943). *Remedial Techniques in Basic School Subjects* (pp. 7-20). New York: McGraw-Hill.
*GANN, E. (1945). *Reading Difficulty and Personality Organisation.* New York: Crown Press.
*GATES, A. I. (1935). *The Improvement of Reading* (pp. 107-117). London: Macmillan.
*GATES, A. I. (1941). 'The role of personality in reading disability'. *J. Genet. Psychol.*, 59, 77-83.
GREGORY, R. E. (1965). 'Unsettledness, maladjustment and reading failure'. *B. J. Educ. Psychol.*, 35, 63-8.
HAMILL, R. C. (1936). 'Emotional factors in mental retardation. A reading problem'. *Arch. Neurol. psychiat.*, 36, 1049-1069.
*JACKSON, J. (1944). 'A survey of psychological, social and environmental difference between advanced and retarded readers'. *J. Genet. Psychol.*, 65, 113-131.
JOHNSON, M. S. (1955). 'A study of diagnostic and remedial procedures in a reading clinic laboratory school'. *J. Educ. Res.*, 48, 565-578.
KNOBLOCK, P. (1965). 'A Rorschach investigation of the reading process'. *J. Exp. Educ.*, 33·3, 277-282.
*LADD, M. R. (1933). *The Relation of Social Economic and Personal Characteristics to Reading Ability.* New York: Columbia Univ. Contrib. to Ed. no. 582.
LANTZ, B. and LIEBES, G. B. (1943). 'A follow-up study of non-readers'. *J. Educ. Res.*, 36, 604-626.
LISS, E. (1935). 'Libidinal fixations as pedagogic determinants'. *Amer. J. Orthopsychiat.*, 5, 126-131.
*MONROE, M. (1932). *Children Who Cannot Read.* Chicago University Press.
*MORRIS, J. (1959). *Reading in the Primary School.* London: Newnes for N.F.E.R.
*NATCHEZ, G. (1959). *Personality Patterns and Oral Reading.* New York University Press.
PECK, L. and MCGOTHLIN, L. E. (1940). 'Children's information and success in first-grade reading'. *J. Educ. Psychol.*, 31, 653-664.
PRESTON, M. J. (1940). 'Reading failure and the child's security'. *Amer. J. Orthorpsychiat.*, 10, 239-252.
ROBINSON, H. A., ed. (1962). 'The underachiever in reading'. *Supp. Ed. Monog.*, 92. Chicago University Press.
*ROBINSON, H. M. (1946). *Why Pupils Fail in Reading* (pp. 76-92). Chicago University Press.
*ROSWELL, F. and NATCHEZ, G. (1964). *Reading Disability* (pp. 63-77). New York: Basic Books.
*SCHONELL, F. J. (1942). *Backwardness in the Basic Subjects.* (pp. 17-49). Edinburgh: Oliver and Boyd.
SOLOMON, R. H. (1953). 'Personality adjustment to reading success and failure'. (pp. 64-82) *Supp. Ed. Monogr.*, 77· Chicago University Press.
*SPACHE, G. (1947). 'Personality patterns of retarded children'. *J. Educ. Res.*, 50, 461-469.
STEWART, R. S. (1950). 'Personality, maladjustment and reading achievement'. *Amer. J. Orthopsychiat.*, 20, 410-415.
STOTT, D. H. (1958). *The Social Adjustment of Children.* University of London Press.
SYLVESTER, E. and KUNST, M. S. (1943). 'Psychodynamic aspects of the reading problem'. *Amer. J. Orthopsychiat.*, 13, 69-76.
*VERNON, M. D. (1958). *Backwardness in Reading* (pp. 132-147). Cambridge University Press.
WOOLF, M. D. and J. A. (1957). *Remedial Reading.* New York: McGraw-Hill.
ZOLKOS, H. H. (1951). 'What research says about emotional factors in reading'. *El. Sch. J.* 512-518.

EFFECTING EDUCATIONAL CHANGE—SOME NOTES ON THE REFORM OF PRIMARY SCHOOL MATHEMATICS TEACHING[*]

by J. D. Williams

National Foundation for Educational Research

Introduction

Appreciable steps have already been taken—especially in America—to reform both curricular content and teaching-method in the teaching of mathematics at the secondary level. Of the many arguments for effecting corresponding changes in the primary school, perhaps the most obvious is that a changed primary education will be needed to prepare the pupil for a changed secondary education. In the present article we shall not venture any further into the rather obscure realms of the justification of reform, but shall confine our attention to matters concerned with the practical problem of implemention, where reform has been deemed desirable. Our observations will not carry with them much of the authority of experience, for very little has been done in the way of systematic investigation of this kind of problem. Neither will they provide anything like a complete solution. We shall briefly suggest some provisional solutions but with no greater ambition than that of helping to define the problem's scope.

Although we shall be concerned mainly with primary-school mathematics, much of our comment will be general in nature, and will be relevant to changes at other levels and in other school subjects.

For the sake of simplicity, we shall consider only those aspects of implementation that fall into the following categories: (i) preparatory steps; (ii) selecting pilot schools; (iii) sustaining the new approach; (iv) teacher-training; (v) evaluation.

Preparatory steps

Where a new approach to teaching has never been tried out, trials in pilot schools will obviously be necessary for its development and for the devising of appropriate teaching-materials. But, even where the approach is not, in this sense, *entirely* new, trials in pilot schools are likely to be needed in any significantly new environment into which the approach is being introduced, so that necessary modifications can be made and so that local teachers can be trained in its use.

The strategy employed in presenting the new approach will depend on the resistances it is likely to encounter in the existing educational system. Where administrators favour the approach, official channels may be used for its introduction, but otherwise a technique of 'infiltration' may be possible —whereby individually-converted teachers foster it until its feasibility and desirability have become manifest to the administration.

---

[*] This article consists partly of material prepared for a conference on mathematics reform in the primary school, held at the U.N.E.S.C.O. Institute for Education, Hamburg, in January, 1966. Its form and content have been greatly influenced by this conference—particularly by the views of a group by which the subject of implementing reform was specifically discussed, the chairman and rapporteur of which was L. G. W. Sealey, Principal of Bletchley College of Education. The proceedings of this conference will shortly be published by the U.N.E.S.C.O. Institute for Education in a report prepared by the present writer.

Mathematics in general—but especially the more mysterious 'modern mathematics'—tends to frighten off the non-mathematician—and primary school teachers are seldom mathematicians. Publicity should therefore be given to the nature of the new curricular content. Entire 'sample' lessons might be prepared for teachers, while booklets might also be provided for the parents of pupils so that they will feel more confident of being able to cope with their children's questions.

Cogent arguments for the proposed changes should be offered. Such arguments might concern: (i) relevance of the new content to modern technological needs; (ii) its importance as a basis for the mathematics to be learnt later in the pupil's career; (iii) its generality, and impingement on other school subjects; (iv) its potential usefulness as a means of training the pupil in basic conceptual skills; (v) the intellectual stimulus it can provide when presented by means of appropriate teaching methods; (vi) evidence derived from objective evaluations, where such evidence exists (see the section on evaluation).

Selection of pilot schools

It is all-important that appropriate schools should be selected for the introduction of a new approach. These should be selected with a view to the possibilities they offer for continued use of the approach, its development, and its evaluation over a sufficiently long time for the results to be of value.

The following features might bear upon a choice of school: (i) selected teachers should have enthusiasm for the approach and for mathematics, and should give some evidence of being able to acquire the new understandings and skills that might be necessary; (ii) more than one teacher per school should qualify in these ways, so that the approach could be continued should the teacher who originally adopted it leave the school; (iii) there should be evidence of staff stability (unfortunately, since teachers who are interested in new approaches tend to be young, promising, or both, they are more likely than others to leave a school on their way up the ladder of promotion); (iv) there should be the possibility of following a cohort of pupils into the later stages of their school career—such a possibility would occur, for example, where pupils from the same infants' school tended to progress to the same junior school, and thence to the same secondary school; (v) the school should be near to an appropriate advisory centre (a training college or university education department, or some other place where expert advice would be available); (vi) the school should be near to other pilot schools, for purposes of administrative convenience, and so that teachers could conveniently intercommunicate or attend meetings; (vii) where pupils are constrained to take traditionally-oriented public examinations, the full adoption of a new approach might constitute a threat to their academic careers; where these examinations are not imposed, or can be modified appropriately, teachers will be less reluctant to abandon a traditional approach; (viii) where schools are larger and multistream in organization, there is a greater likelihood that participating teachers will be allowed extra time for work on the new approach; (ix) apparatus and group work involved in some of the modern teaching techniques require a certain amount of space—where accommodation is generous, such techniques stand a better chance of being used properly; (x) naturally, where the local educational authority smiles upon the new approach, a school is likely to receive more encouragement and help than otherwise.

Sustaining the new approach

As has been pointed out elsewhere[*], it is all too easy for a teacher who has adopted a new approach to revert to a style of teaching that is more appropriate to approaches to which he is more accustomed. Measures should be taken to ensure that such 'method-reversion' does not occur. Such measures might include the following: (i) at the stage of inservice-training the basic teaching 'philosophies' underlying the approach should be communicated to the teacher; (ii) he should be given individual responsibility within the general scheme of the reform, so that his personal involvement in the approach will be increased; (iii) communication among participating teachers should be facilitated, so that social support and stimulus are provided; (iv) initially, in order that it should enjoy a good chance of success, the approach should be tried out on bright pupils in small classes; (v) instruments or procedures should be provided for evaluating pupils' suitability for various stages of the new curriculum; (vi) instruments of evaluation that are appropriate to the new teaching content should be provided, so that the teacher can judge his pupils' progress by relevant standards, and thus carry out a concurrent evaluation for himself; (vii) within the broad aims of the new approach the teacher should be allowed a certain amount of freedom to select mathematical topics, methods of developing these topics, and teaching materials; in this way he will be able to cater for his personal preferences and to diagnose his personal strengths and weaknesses in relation to the teaching of the new content; (viii) free time should be allowed for the preparation of lessons; (ix) specialists should be available for consultation at all times—even if only by telephone; (x) an 'assessor' might regularly appraise the teaching situation, and offer advice in the light of such features of the situation as topics being taught, levels of achievement and ability, methods of presentation being used, the teacher's attempts at sequencing and interrelating parts of the curriculum, the grouping of pupils, the length of lessons, and the bearing of such factors as these on the teacher's objectives.

Systematic concurrent assessment, like that mentioned above, is primarily for the teacher's own guidance, but, since it involves the use of standardized procedures, the information obtained from different schools by its means will be amenable to collation in various ways, so that some kind of general picture of the effect of the new approach may be built. However, such assessment—involving, as it does, the subjective opinions of teacher and assessor—cannot take the place of a rigorously objective evaluation in its role of arbiter of whether the new approach should be employed on a large scale.

Teacher training

The material included in this section draws heavily on the report: 'Mathematics Teaching in the Primary School', compiled by Z. P. Dienes[†].

At the point at which the beginnings of a reform are taking place it will be necessary to look to all kinds of sources for help in the inservice training of teachers. Dienes and his collaborators suggest, among other possible sources, the following: (i) experts in subject-matter might be drawn from local training colleges or universities, or even from among pupils' parents; their function could vary considerably—from giving teachers formal lectures to helping them in the classroom; (ii) local study-groups of teachers could be formed, which might or might not include a specialist consultant—if not, written sources of information and ideas for discussion might suffice; (iii) novice teachers could visit classes in schools other than their own, in which the new approach had already become established; (iv) courses could be run by university departments or teacher organizations; such courses might range from purely expository lectures to working sessions; (v) specialist teachers within each school might deal with the task of mediating changes in curriculum and teaching method; (vi) films, television, and radio could be used—especially for communicating to teachers the subtler details of teaching technique; (vii) self-study materials, such as conventional textbooks, programmed textbooks and film strips, are on the market and could be used; journals on primary-school mathematics teaching might also stimulate and enlighten; (viii) the recording or filming of teaching sequences permits their subsequent analysis, which can be most instructive—especially to the teacher whose behaviour is being analysed; (ix) apart from its value as a means of making highly effective use of available teaching expertise, *team teaching* might enable teachers both to learn from one another's example and to improve one another's technique by a process of mutual analysis and criticism; (x) as a means of providing incentives for inservice training, formal qualifications could be awarded to those who had completed certain kinds of inservice course.

---

[*] *See* WILLIAMS, J. D. 'Method-reversion: the problem of sustaining changes in teacher behaviour'. *Educ Res.*, Vol. VIII, No. 2, February, 1966.

[†] This is due to be published soon by the U.N.E.S.C.O. Institute for Education, Hamburg.

Evaluation

The question of whether or not the aims behind a new approach are proper is a matter for *argument*—but there are many questions which, once they have been formulated in operational terms, are amenable to empirical investigation. Given that certain aims are acceptable, does the new approach achieve them? Are these aims achieved more effectively by one particular approach than by others? Given that these aims are achieved best by a particular approach, what is the effect of the achievement of these aims in relation to broader aims? For example, what side-effects on the teaching of other subjects accompany the introduction of a new approach?

Too often reform movements pay more regard to fashion than to fact. Very often objective evidence bearing upon the feasibility of reform and the directions it should take could be obtained by a variety of reliable research procedures, but very seldom is this done. The confidence with which reformers impose their hunches on the educational system without so much as a gesture towards the assessment of outcome must seem strange to those who work in fields where product- and process-testing are established routines.

We shall end these notes with some brief comments on the difficulties involved in one kind of evaluation of new approaches, together with some suggestions indicating how these difficulties might be overcome.

From the point of view of deciding to change from one teaching-approach to another, the most pertinent kind of evidence would come from a straightforward comparison of the effects of adopting these alternative approaches. However, the difficulties in making such comparisons are immense. Many of these are analysed in an earlier article*, so we

---

*See WILLIAMS, J. D. 'Some problems involved in the experimental comparison of teaching methods'. *Educ Res.*, Vol. VIII, No. 1, November, 1965.

shall consider here only a few—and these only briefly. (i) At which stage in the child's career should an attempt be made to assess the effects of the teaching to which he has been subjected? Where, as in the case of most new approaches to mathematics teaching, we are claiming a long-term payoff (expressed in such statements as 'laying the foundations for later learning', 'providing basic understandings rather than specific skills', etc.) this problem becomes acute. (ii) How can we compare the outcomes of approaches whose teaching aims are different? Direct testing of the relative effectiveness of approaches that share very little common curricular content begins to present us with the kind of problem that we should meet in trying to compare colour of hair with length of nose. (iii) In the early stages of the use of a new approach we cannot be sure that its effect will be the same as it will when the approach has become established. On the one hand, until those using the approach have become experienced in its use, they may not have the skill to do it justice. On the other hand, the very novelty of the approach may stimulate all concerned with it to do it *more* than justice. In addition to these two difficulties we have to face the dilemma posed by the fact that if we encourage the initial establishment of a new approach by making conditions favourable, we at the same time give it an unfair advantage from the point of view of assessing it.

Various ways of coping with these, and other difficulties, have been suggested. For example, learning experiences have some immediate effects which we could expect to influence later learning; by assessing such effects we might obtain information that would bear upon later stages in the pupil's career. Of these, we might measure the understanding of basic concepts, or, by utilizing as tests certain standardized programmed instructional sequences, we might measure preparedness to learn certain topics. Again, to some extent the aims of different approaches may overlap—for example, both approaches may aim to teach the structure of the decimal number-system or to leave the pupil with a liking for mathematics. In such areas of overlapping aims, a common measure is possible. Again, it could be said that there are some accomplishments that we demand of the pupil whether or not an approach *aims* to produce them. Skills at applying mathematics in real-life situations might be included among these. An approach that did not produce such accomplishments could be condemned, *whatever* its aims might be. Finally, if we could prove that the new approach was equal to its rival in achieving its rival's objectives, we might vindicate it at least in the eyes of those who ask only that no positive harm should derive from its adoption.

Whatever the means of evaluation chosen, it is vital that appropriate steps should be taken *at the outset* of a reform project to ensure that objective and independent evaluation of the kind discussed above will be possible. Such steps might include: the precise definition of the reformer's aims in ways amenable to objective study; the selection of a sample of schools into which to introduce the new approach, and which will not prejudice outcomes in ways that cannot be allowed for; the pretesting of participating pupils; the testing of 'control' pupils who attend the same schools as these but who belong to a generation for which the traditional approach has been used. Even where the final decisions about evaluation have not been made, it may be necessary, before introducing the new approach, to take such steps as these in order to ensure that the doors to certain kinds of investigation remain open.

SCHOOL REPORTS

by R. E. Copland

Deputy Headmaster, Wyndham School, Egremont

CONSIDERABLE research into the compilation and use of school record cards has been undertaken (see, for example: Schonell, 1948; Walker, 1955; Peel, 1958) but very little into school reports intended to be read primarily by parents.

This article gives a personal view from the standpoint of a secondary school teacher. After listing aims and criteria a typical school report is analysed and a number of serious imperfections noted. Some existing practice and literature are then reviewed and the development of a modified procedure and format is described. Finally, possible future lines of inquiry are outlined which, it is hoped, may encourage further research on a topic of widespread interest.

Aims and criteria

It would probably be generally agreed that the aims of a report include the following:

A. To give parents a clear indication of the pupil's progress in school work with respect to his year group as a whole, his form or set and his own normal standard.
B. To give parents some indication of his behaviour and attitude to school work, his personality and growth in maturity.
C. To encourage the pupil to develop his abilities to the full.

Further criteria which could be applied in judging a report are:

1. Is it attractive in appearance?
2. Are the staff able to fill in each section of the report in a manner which is obviously meaningful, relevant and unambiguous to them?
3. Is the report efficient in terms of useful information given for energy expended in compiling it?

The need for an attractive report in these days where design and layout of leaflets, books, etc. are of such a high standard seems self evident. The fulfilment of the second criterion will mean that reports can be compiled with care and interest. In the writer's experience, staff are understandably reluctant to fill in information sheets—particularly record cards—in which there appears to be a large 'hit and miss' element. For example, a teacher taking only, say, form 3C in the third year and asked to assess each pupil for sociability as –, normal, or + (20%, 60%, 20%) over the year group as a whole is in considerable difficulty to say the least. Although he may be their form teacher he sees them in very restricted circumstances. The third criterion is included to try to prevent too high a proportion of the staff's energy for a week or two being absorbed in report writing. The activity frequently affects marking and preparation and, in extreme cases, can bring the life of the school practically to a halt.

Acknowledgement:

The writer wishes to express his appreciation of the encouragement given by the Headmaster, Mr. J. R. I. Sharp, and staff of Wyndham School in the preparation of the loose-leaf report described in this article.

A typical school report

Reports are normally either on single sheets, approximately foolscap size, or in bound book form with a page for each report. The advantage of a bound book is that the school has readily to hand a record of the pupil's progress and so can easily refer to previous remarks when compiling reports. Disadvantages are: first, that parents cannot keep reports until the book is complete; and second, that a school's decision to use this format necessarily involves commitment to it for a considerable time ahead. With separate sheets each report can be retained by the parent and new formats can easily be made; but since many schools do not make and retain copies, it frequently happens that each report is an entity in itself rather than the progressive development from previous ones which it should ideally be.

How do school reports fulfil our aims and criteria? The majority, in either sheet or bound book form, are basically of the format shown with a number of common variations, e.g. an unstandardized grade rather than a percentage may be given; there may be separate columns for examination and term marks; a column for assessment of effort may be included—but these will not greatly affect the following discussion and conclusions.

Name.................................... Term .... No. of absences..........................

Form .... No. in Form...................... Final Position..............................

Conduct..............................

SUBJECT	%	POSITION	REMARKS	INITIAL
English				
Art				

School activities

Form Teacher's Comments

Headmaster.................................... Parent's signature....................................

We shall consider each section of the report in turn:

Form

Often no indication is given of the ability range covered by the form or set. This may not have mattered when schools were usually either two-stream or three-stream labelled A, B, and C; but now, with large schools and multifarious ways of naming forms, confusion easily arises.

Final Position

The implication here is that to add together results of performance in English, Art and all other subjects leads to a meaningful assessment. It seems reminiscent of those first lessons in algebra in which one was taught that to add four apples to five nails resulted in ' nine ' but little else. Whilst use is often made of final positions for promotions and demotions, frequently some such device as extra weighting for English and Mathematics or totalling only ' academic ' subjects is resorted to. These devices acknowledge the weakness of the concept of final position but at the same time make no allowance for the relative spread of the marks (i.e. the standard deviation) in each subject—the only factor determining the influence of the marks of a particular subject on the final rank order. It is a popular misconception that the set of marks with the highest average has the greatest influence but the following example exposes this fallacy:

Pupil	*Maths Mark*	*English Mark*	*Total*	*Rank*
A	30	90	120	3
B	40	85	125	2
C	50	80	130	1

The rank order is determined solely by the mathematics marks with the greater spread. Incidentally, it is rarely noticed in staff rooms that the standard deviation of mathematics marks (where percentages of 0 and 100 are quite possible) is frequently greater than that of English, where 25 or 30 will often be a minimum and 70 and 75 a maximum!

Conduct

Comments are frequently limited to such remarks as excellent, very good, satisfactory and unsatisfactory as if these comprehensive and static terms applied to the whole range of the pupil's behaviour and conduct at all times.

Percentage

Without further information, the figures in this column are, of course, meaningless, the percentages obtained being quite as dependent on the difficulty of the tests or examinations on which they are based as on the performances of the pupils. 60% in English may frequently represent an inferior performance relative to the class as a whole compared with 50% in, say, Mathematics.

Position

The rank order certainly gives more information than the percentage figure, although it gives no indication that, in a normal distribution, the probability is that differences between ranks 1 and 2 or 29 and 30 (in a class of 30) will be much greater than those between, say, 15 and 16. It does not necessarily follow that equal ranks in different subjects represent equivalent achievements, and at the lower end of the scale an unduly optimistic picture may be given if the number being assessed is less than 30 due to absence or a small option group. Nevertheless, to me, the main objection to the rank order is the implication of a race or game in which the prime purpose of learning is to beat one's neighbour.

Remarks

As the reports have to be passed from teacher to teacher, this often leads to their being filled in hurriedly or at an inconvenient time. On most sheets, insufficient space is given for the occasional pupil about whom one would like to write in some detail.

The comments, however, may be constructive and directly concerned with the teacher/pupil learning situation, e.g. ' Achieves a high standard of practical work but slowness in writing is a handicap '. On the other hand the remark may be a blunt one such as ' Idle ' or ' Unable to concentrate '. This may genuinely reflect the pupil's behaviour in a particular subject, *as observed by his teacher*, though as the comment is not constructive it does not encourage the pupil. Moreover, words and phrases like these are emotive and likely to be widely interpreted by busy headmasters, other staff, parents and the pupil himself so that in effect they become personality trait assessments. Now, as personality traits certainly have an important, if not always clearly defined, bearing on school performance, one might assume that there is no harm in this. But such assessments are made under very unsatisfactory conditions with few, if any, of the most elementary rules for personality assessment being observed.

Form Teacher's Remarks

As with subject remarks these are closely connected with personality assessment and are highly subjective so that the same pupil can receive very different assessments from consecutive form teachers. Nevertheless this may largely reflect a real difference in the teacher/pupil relationship and the remarks can often be useful providing that they are confined to situations in which the teacher has personal knowledge of the pupil's behaviour and work.

Parent's Signature

A signature alone is inadequate and easily forged! Where an opportunity for written comment by the parent is provided this is often valuable but more important still is the provision of an early opportunity for the parent to consult with the school staff, with the report before them. Efficiently arranged, such a meeting can be one of the best ways of enabling the report to fulfil its basic aims. It acknowledges that, in any complex field, communication, to be effective, must be two-way.

Apart from specific inadequacies in reports, a number of general aspects seem unsatisfactory. These include the following:

1. Conditions for the halo and stereotype effects are as favourable as they could be. Not only can each teacher see what his colleagues have written—and which of us is not tempted to strengthen or weaken the remark we intended to make if it is totally different from others appearing on the form?—but often a pupil's report is written in the staff room to a background of comment and discussion about him.
2. Too rarely are reports looked forward to by reasonably seriously-minded pupils as a help to their studies and development.
3. The average report is distinguished by its unaesthetic dreariness.

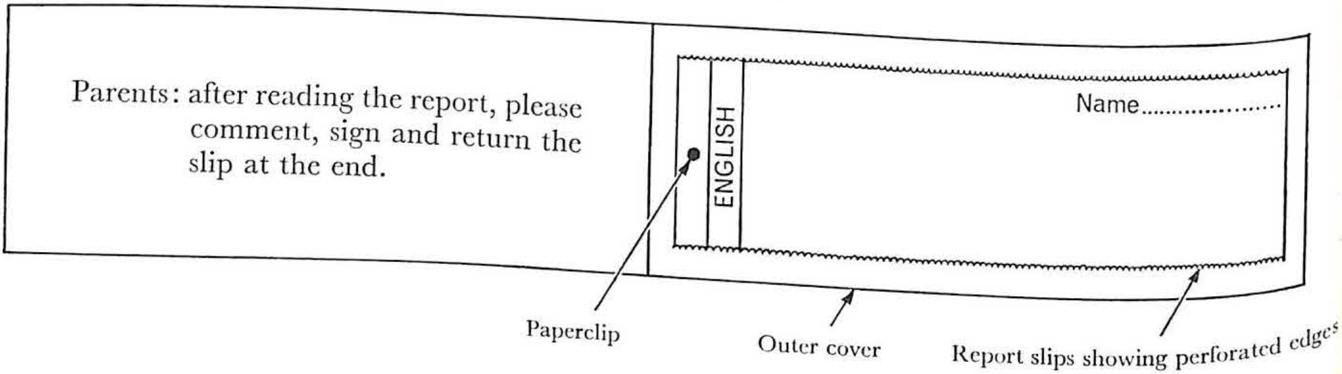
Existing practice and literature

In this section a variation on the typical school report is considered together with some literature relevant to possible developments.

Physical Form

An interesting alternative to the sheet or bound book is one understood to have been developed between the wars at Eton College and known to be used currently at Fettes College and Taunton School. Here the report is a booklet consisting of an outer cover approximately 8″ × 3″ with a separate slip for each subject, form teacher's remarks, etc. inserted and held together with a paperclip.

Booklet type of report (**opened out at first page**)


Parents: after reading the report, please comment, sign and return the slip at the end.
ENGLISH
Name..................
Paperclip
Outer cover
Report slips showing perforated edges

Subject reports come to teachers as perforated foolscap sheets each containing five report slips which can be filled in at leisure. Carbon copies can be made if required, and the halo effect between teachers is greatly reduced. Furthermore there is plenty of space for each subject, printing can be adapted to the requirements of any particular subject and changes in format can be made cheaply. The cost of about 3½d. a report per pupil, or 4½d. if a copy, without cover, is to be retained by the school, is relatively high but scarcely excessive if the information it contains is worthwhile. There is also some clumsiness for filing purposes and, as only one page at a time is visible, inefficiency in reference. For junior forms there is the added disadvantage that, as for aesthetic reasons the booklet must have certain minimum dimensions of the order of 8″ × 3″, this leads to rather excessive space if the briefest of comments only is required. Despite these minor disadvantages, the booklet type of report seems worth serious consideration particularly for senior forms.

Format

Pattinson (1957) describes an experiment in a secondary modern school in which the report seems to have been rather like a certificate. Subjects were listed under two headings: those in which good work was being done and those in which exceptionally good work was being done, subjects not in either category being omitted. Comments, if made at all, had to be favourable and/or positive. Examples given of acceptable comments include: excellent work, good progress made, is obviously trying much harder, always makes good oral contributions in discussion. Whilst there will be many reservations concerning the limitations of a report of this type, Pattinson claims that it was popular with the pupils and improved their attitude to work. Incidentally the reports were prepared in loose leaf form and they were bound into books in the Art and Craft department.

Rowe (1959) reproduces a report form he has used in which a progressive type of remark has been introduced. Thus under the heading, 'Growth as a Person: 1. Behaviour', each pupil is assessed as: always/mostly/sometimes/occasionally/not yet, behaving responsibly and considerately. Under, 'Attitude to Work', each pupil is assessed against each subject as, A: has developed the power and will to work on his own; B: is beginning to develop the power and will to work on his own; or C: shows *as yet* (my italics) little or no sign of developing the power and will to work on his own. Once one has noted the implication of hope in progressive rather than static remarks the importance of this element in the aim of encouraging the pupil is clear.

Rowe (1964) gives a striking illustration suggesting that the ratio of handwriting to print is often an indication of the value of a report. No doubt it is, but an interesting contrast is provided by the University of Texas (1957) who sponsored a comprehensive investigation into the methods of reporting to parents in America where extensive use of check lists is made. In the so called SIU system, each subject may have up to a dozen comments to be checked under columns headed: Satisfactory, Improving, and Unsatisfactory. A system such as this certainly gives the parent a comprehensive assessment, and more use of check lists could possibly be made in this country. Nevertheless there are considerable difficulties in covering all aspects of the pupil's performance in this way especially if a 'progressive' element, as exemplified in Rowe's report, is to be included.

Assessment of Attainment

The difficulty of interpreting percentages and the possible harmful effects of rank orders have already been mentioned. Standardization of marks is, of course, an essential procedure for any serious study involving comparisons yet it is rarely done for school reports. A mean of 50 and a standard deviation of 10 gives small numbers for calculation and the scores are easily compared, in nearly normal distributions, with T scores which are based on rank order (see for example: Garrett, 1958). However, Daniels (1949) recommends a mean of 50 and a standard deviation of 15 as most suitable for school use because the scores do not look very different from the familiar percentages. This could be an important consideration in the early stages of introducing standard scores, especially as very little extra computation is needed. Copland (1962, 1963) has shown that extensive standardization of marks in school is a practical possibility and, with suitable organization including the use of an adding machine, involves no more work for individual teachers than a system of percentages.

Personality Assessment

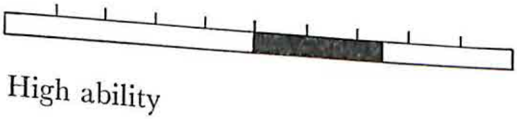
A comprehensive review of personality traits and their relationship to school attainment has been given by Warburton (1961). A study of the section, 'Rating Techniques', shows immediately how unsatisfactory are the formats and procedures of school reports for the sweeping assessments of personality which are often made. In fact, even under the most stringent conditions, the assessment of personality is fraught with difficulty (Vernon, 1964). If however, satisfactory conditions can be produced, the proposals of Warburton (1962) are interesting. He suggested that useful assessments for report purposes would be persistence, intellectual interests and independence of judgement. These are based on research findings which suggest that Cattell's factors G, F– and Q_3* are the most important in scholastic attainment.

Towards the development of more satisfactory school reports

We shall now discuss how to overcome the specific objections that the report gives incomplete, misleading or uninterpretable information.

The range covered by the *form* or set can be indicated by listing the approximate ability percentiles covered by the form (e.g. 50 – 75 for a C form in a 4–form year) or perhaps, better still, by shading in a grid as shown below:

Form 3C. Range:



The statement of *final position* seems best dealt with by omitting it, whilst *percentages* and *positions* can be replaced by standard scores. Would parents, however, be able to understand a system of standardized marks? There might be some initial confusion, but there seems little difficulty in the basic data of the scale, namely that the mean for the form is 50, approximately two thirds of the pupils will obtain scores between 35 and 65, and scores of 75 and above or 25 and below indicate that the pupil is respectively substantially above or below the average of the form. This certainly seems clearer than the alternative of: ' Spring term mathematics—60%, Summer term mathematics—50%,

---

*For a description of these traits, see for example, Cattell, 1957, 1965.

Comment: his work has improved ' (which may well be true). Not all teachers will wish to allot marks to work done, and in cases where grades are allocated these too can be standardized by similar procedures.

The writers of comments under the headings of *Conduct*, *Remarks* and *Form Teacher's Remarks*, rather than giving overriding consideration to their function as a record of past impressions, should give due weight to their personality assessment implications and to their influence on pupils' future attitudes to study. The aims of the report and the value of constructive comments which leave room for hope, need stressing to school staffs more frequently and forcefully. It would also, no doubt, be wise to warn against the danger of extrapolating observed behaviour in restricted circumstances into wider generalizations, and to give examples of static and unconstructive remarks.

It was suggested earlier that *parents* should have an opportunity of discussing reports with the staff. The fact that we may need to justify our comments face to face is, for many of us, a desirable constraint during compilation of the report. The effectiveness of a meeting of this type is greatly enhanced if it takes place within a few days of the receipt of the report. On such an occasion, it is most desirable to have tables for each member of staff labelled with his name and subject, and an invaluable associated ' visual aid ' is a prominently displayed blackboard with staff names and signatures!

There remains the question as to whether any specific personality traits should be formally assessed on the report. The following arguments concerning persistence, already appearing on many reports as ' effort ', will apply also to other traits. For internal school records there may often be a strong case for assessing persistence despite the practical difficulties*. Undoubtedly the results of such an assessment would be of value when preparing entries for the ' remarks ' columns of reports, but there would appear to be good reasons why an assessment of persistence should not appear as a grade or number on the report form. If, as is frequently the case, persistence or effort is to be directly assessed in a column adjacent to attainment, the correlation between them is likely to be high because of the halo effect. Even if a persistence score is given following a strict procedure, there may be difficulty in explaining to parents and pupils the basis of the assessment. Although, in any case, this is subjective, it might be interpreted, because it appears as a letter or number, as in some way specially authoritative—and more so than a corresponding comment in the remarks column. If this happens, the effect of rating a person highly for persistence when he, himself, knows that this is not true could seriously affect his future attitude to work. Even more serious would be the effect of an error in the other direction.

Modification of the report sheet and book

A report prepared with the above considerations in mind should be a more informative and helpful instrument of communication between school and home. We shall now consider how it can be made as attractive, flexible and efficient as possible.

Where a simple report is required a marked improvement in appearance on the normal foolscap sheet can be made for small extra cost by using a cartridge weight paper about 9½″ × 9″ folded to give a 9½″ × 4½″ booklet, printed in colour and with attention given to the typography. The reduced size is compensated for by being able to use all four pages and, of course, a variety of layouts is possible. One that has been used is as follows:

---

*These are described, and a possible procedure for assessment is outlined in the Appendix to this article.

Page 1. Name of school, name of pupil, house, form and date.

Page 2. Subject performance.

Page 3. Special activities and interests, Housemaster's and Headmaster's comments.

Page 4. Notes on grading system, forms, sets and house groups.

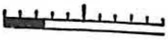
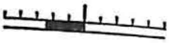
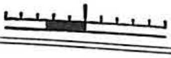
However, it may often be felt that a more permanent record is needed. Following consideration of the principles already discussed, a loose leaf book has recently been prepared and is to be used for all pupils below the sixth form.

The cover selected is a 'Boa' flat quarto size spring back file made to order without its standard trade mark printing. A name label in a contrasting colour has been added.

Each page contains a record of the pupil's work over two or three years in one subject and could have the format shown, though that chosen by the writer's school differs from this in some respects:

English

Name..............................

Year	*Term*	*Form or Set*	*Attainment (Standard Scores)*	*Remarks*	*Initials*
1965/6	*Autumn*	**I.A** 	**30**	**Tries hard but finds the work difficult.**	**J.S.**
	Spring	**I.B** 	**55**	**Has settled in well.**	**T.J.**
	Summer	**I.B** 	**58**		**T.J.**

Note: The **bold** type indicates what would be written by the teacher.

Information for the guidance of readers on interpreting attainment scores and the significance of the range scales, can be printed on the inside cover. A printed sheet can carry the Form Teacher's (or Tutor's or Housemaster's) and Headmaster's remarks, with additional sheets if required for activities and parents' comments.

Year 1965/66 Name........

AUTUMN TERM	Form Teacher Headmaster

Conduct and attendance could easily be indicated though it seems reasonable to assume that these are unexceptional unless commented on by the form teacher.

It is hoped that the 'page to a subject' rather than 'page to a term' system will have the following advantages:

1. The file can be quickly sorted into subjects so that each teacher can write his reports in his own time. (The form teacher's and head's comments appear on the same sheet as they are normally written consecutively.)
2. There is no halo effect between subjects.
3. Previous records in a given subject are immediately available to the teacher but clearly this creates the conditions for halo effects within a subject. Indeed, standard scores may themselves accentuate the conditions but information concerning previous performance will generally be regarded as most desirable.
4. Varying space can be allowed on the record for different years.
5. On any report, it may be felt that a comment is not always necessary but this may lead to some pupils getting very few comments at all. Here a minimum quota of, say, one comment per subject per year, could be agreed.

6. Increased responsibility would be placed on each department and teacher and could help towards more constructive comments over a period of time. Thus a comment, 'Shows no interest in the subject', if repeated too frequently is very obvious to the teacher responsible for the overall progress of the pupil, and could well prompt an inquiry as to what the department was doing about it.

Naturally, a progress record such as this is a compromise, and obvious disadvantages are cost (estimated at 3/6d. for the file and 15 sheets of ledger quality paper), the effect on the insert sheets of the handling involved and, from the form teacher's and head's point of view, the extra time needed to read through the record each term. Nevertheless, it is hoped that the progress record will go some way to satisfying our aims and criteria, but the quality of the remarks will remain a key factor.

Conclusion

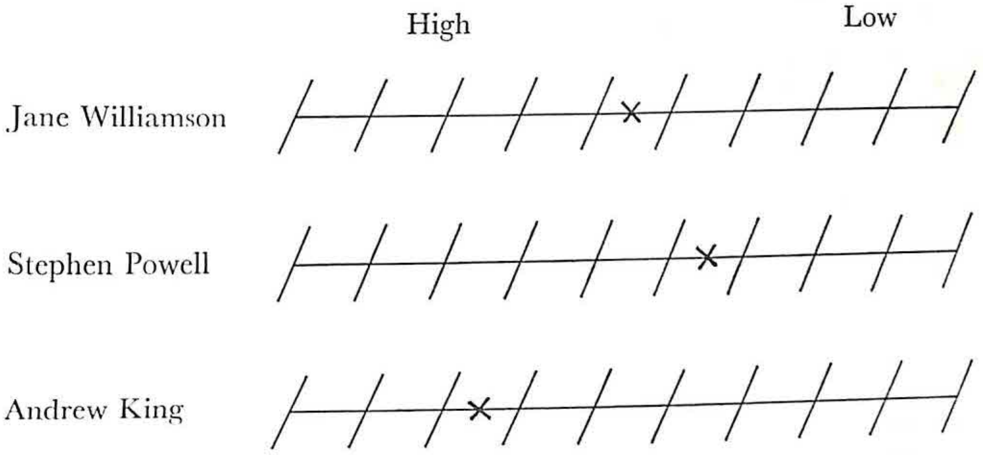
Many possible topics for research, such as the degree to which 'remarks' columns could and should be made more objective, and the appropriateness of various type situations for personality trait assessments at different stages of school life, have been implied or suggested. Other questions could be asked, for example: are pupils and parents depressed by being continually reminded of past failures if the book or file contains previous reports? How extensively and by what means should reporting back by parents and pupils take place? Would it be advantageous to invite pupils to submit self-reports? The need for more research into the whole field of communication between school and home is clear but, from the schools' point of view, perhaps the first priority should be investigation into the attitudes of parents, pupils and teachers to various forms and details of school reports. This would provide a firm foundation for future progress.

Appendix. Assessing persistence

There appear to be three main sources of difficulty in assessing persistence in school. Firstly, most teachers, particularly in a large school, will have direct knowledge of isolated forms rather than year groups. In a given class, especially if it is a keen one, the range in persistence may be very narrow. One could, perhaps, rely on past knowledge of the range of persistence in the year in question, but this seems a most unsatisfactory procedure. Secondly, the situations in which each pupil is directly observed by the assessor are limited in scope for persistence, especially in more senior forms or in subjects with a small time allocation. In these cases the most important manifestation of persistence may take place during homework or at other times outside the teaching period. Finally, in an experiment tried by the writer in which an attempt was made to assess the persistence of 16 fifth-year pupils to whom he had taught mathematics for three years, as great a difficulty as any was to detect unsuccessful persistence especially if the person under consideration was not one who asked many questions in class.

Despite these difficulties, experiment and experience are important and a suitable procedure for assessment might be as follows:

1. List type situations likely to be encountered by the teacher. This will help to define persistence for the purposes of the assessment and may well lead to a narrower interpretation than usual, e.g. persistence as observed in the classroom.
2. Prepare form lists in a scrambled order to reduce halo effects between attainment and persistence. If an alphabetical list is used in a mark book, the teacher may know only too well that the weak James Logan follows the able Andrew King and this may inadvertently weight his assessment of persistence in the same direction.
3. Fairly quickly assess each pupil by placing a cross along a segmented line



4. Total the number of crosses in each section

		1	2	5	13	6	3	

The most suitable method of presenting the results will depend on the time available and rigour required. A very simple approach would be:

Jane Williamson 3 ⑤ 22

indicating that 3 pupils rated better than Jane who was in a group of 5 followed by 22 with lower assessments. Alternatively tables for different sizes of class could easily be prepared to convert the raw totals into percentages. The figures would then be:

Jane Williamson 10 ⑰ 73

It would be much better, of couse, to convert the percentages into T-scores or standard scores and this would not be difficult with the use of suitable tables.

References

Cattell, R. B. (1957). *Personality and Motivation Structure and Measurement.* London: Harrap.
Cattell, R. B. (1965). *The Scientific Analysis of Personality.* Harmondsworth: Penguin.
Copland, R. E. (1962). 'Testing, marking and assessing'. *Mathematics Teaching,* No. 20.
Copland, R. E. (1963). 'Letter to the Editor', *Mathematics Teaching,* No. 23.
Daniels, J. C. (1949). *Teachers Handbook of Test Construction, Marking and Records.* London: Crosby Lockwood.
Garrett, H. E. (1958). *Statistics in Psychology and Education.* New York: Longmans.
Pattinson, W. (1957). 'A Study of the Changes brought about in the Educational and Social Climates of a Secondary Modern School over a period of five years'. Unpublished M.A. thesis, University of London.
Peel, E. A. (1958). *The Psychological Basis of Personality.* Edinburgh: Oliver & Boyd
Rowe, A. W. (1959). *The Education of the Average Child.* London: Harrap.
Rowe, A. W. (1964). 'School Reports', *Where,* No. 17.
Schonell, F. J. (1948). *Backwardness in the Basic Subjects.* Edinburgh: Oliver & Boyd.
Texas, University (1957). *Four Methods of Reporting to Parents.* University of Texas.
Vernon, P. E. (1964). *Personality Assessment.* London: Methuen.
Walker, A. (1955). *'Pupils' School Records: A Survey of the Nature and Use of Cumulative School Records in England and Wales'.* N.F.E.R. Publication No. 8. London: Newnes, for the N.F.E.R.
Warburton, F. W. (1961). 'The measurement of personality'. *Educational Research,* Vol. 4.
Warburton, F. W. (1962). Private communication.

THE TESTING OF SPOKEN ENGLISH IN SCHOOLS AND COLLEGES IN THE UNITED KINGDOM

by P. J. Hitchman

Department of Education, University of Nottingham

RESEARCH studies dealing with the testing of spoken English were until recently all of American origin. They were of course designed to serve American education.* The institution of the London 'O Level' tests and the new C.S.E. tests in Oral English has acted as a spur to British investigators, and studies originating in the United Kingdom are now appearing. Available to date are those of Hitchman (1963, 1964, 1966b), The English Speaking Board (1964), Wilkinson *et al.* (1965) and the Southampton Experiment (1966).

Although there has been some overlap, each study has had a distinctive theme. Hitchman was concerned to demonstrate the degrees of validity and reliability of tests currently on offer (each set of candidates being tested twice with 'equal and parallel' test material). The English Speaking Board was not concerned with 'external' validity, and tested its subjects once only to gauge the amount of agreement between a *trained examiner* and a group of *untrained moderators*. Wilkinson and his associates brilliantly broke new ground with tests of Listening Comprehension; while the Southampton Experiment was designed to compare four commonly suggested methods for testing oral expression—(1) Reading a passage of prose and talking with the examiner. (2) Making a short speech or lecture and answering questions. (3) Talking to the examiner about a diagram previously studied. (4) Participating in group discussion.

The studies reveal a number of interesting features, and a brief summary of these, with some comment, may be useful.

(1) *The Testing of Listening Comprehension*

Wilkinson and his associates think that teachers should deal with the real facts of the English language. How do people *really* speak? The new linguistic approach to the study of English usage enables researchers to examine this question under three headings. The first divides people into regional groups, that of Dialect and Accent; the second into professional or occupational or social groups, that of *Register;* the third is the division by the type of situation in which speech takes place, and the appropriateness of the discourse, that of *Style*. 'English is there, in use all the time, by different groups, in a variety of situations. Let us show children what there is, and showing them, help them to a wider control and a greater tolerance of other registers, other styles.' Using the techniques of linguistics the authors have devised a test in which children, listening to a good tape-recording, are asked to attend carefully to such aspects of a speaking situation as *Content*,

---

*See an article by the present writer in *Educational Research*, Vol. VII, No. 1, November 1964.

Register and *Style*, and *Phonology*.* Validity and reliability are guaranteed by standardization and completely objective (multiple-choice) marking. Samples of the test material devised by Wilkinson and his associates are so interesting that one is afraid of a rush of tapes and booklets consisting entirely of similar material from which all teaching of Listening Comprehension is to be done. The authors say that the 'backwash' of Listening Comprehension tests will be in terms of increased *awareness* of the spoken language and how it is used in differing circumstances.

(2) *Tape-Recording Assessment*

The possibility of assessing spoken English by tape-recording is tantalizing researchers and test authorities. Some authorities and other interested parties will have nothing to do with it—their argument being that so much of feeling and meaning is carried by the manifestations of bodily presence that no assessment can be reliable which is based on evidence that does not include these manifestations. However, this is yet to be proved. For some years the Brighton L.E.A. has used tape-recordings as evidence additional to face-to-face testing, and this is now the standard practice of the South East Regional C.S.E. Board in its oral English testing—the tapes are used to 'moderate' the 'live' results. And, as we have noted, tape-recordings were used for Wilkinson's Listening Comprehension experiment.

The first evidence we have as to the validity and reliability of tape-testing comes from the Southampton Experiment. In each of the four types of test noted above an examiner conducted the test, and he and a non-participating observer assessed the performances. In the first three of these tests (i.e. in all but the Group Discussion) tapes were made, and these were assessed quite separately by an independent marker who had not been present at the actual testing. The experimenters claim that assessment by tape-recorder is a practical proposition, although the coefficients of reliability between examiner and observer were higher than those between examiner and tape-assessor and observer and tape-assessor (r being about 0·8 on the one hand and varying between about 0·7 and 0·6 on the other). However, they report that recordings in some cases were impaired by heavy interference due either to technical flaws or a noisy site.**

(3) *Group Discussion*

This was the fourth test in the Southampton Experiment. Groups were five or six strong and a session lasted about 25 minutes. Candidates were given three topics half-an-hour before the test; the subjects were used as starting points and the conversation ranged wherever the candidates took it. A reliability coefficient of 0·785 was found for examiner v. observer and validity coefficients (correlations with teacher's assessments) of 0·344 (examiner) and 0·41 (observer). The experimenters deduce that Group Discussion measures ability unrevealed in 'the normal classroom situation'.

---

*Phonology, as used here, means the study of the sound-structure, or system of sounds, in present-day English. Thus children can be asked to attend carefully to sound contrasts (e.g. pronunciation contrasts) found in diverse samples of spoken English and, as the author of this chapter of the report (Davies) points out, also to grammar contrasts (as in Intonation, which clearly acts as a grammatical system of language just as much as tense, syntax and so on).

**Obviously, technical experimentation is required. What is needed is a means of making an easily 'readable' recording in *any ordinary school classroom*, one that is within the means of any secondary school. Such recordings have been made in a day's experimentation in a secondary modern school under very trying noise conditions by the writer who used a well-known make of tape-recorder, a modern unidirectional microphone and a lightweight three-fold acoustic screen 6′ × 4′ with a roof 5′ × 4′, made of expanded polystyrene.

(4) *Prose Reading and Conversation*

This test-combination is widely popular with test authorities. The Southampton experimenters considered this type of test 'rather disappointing' and the least satisfactory of their four test methods. Separate marks for Reading and Conversation were not given. The examiner and observer produced an agreed mark and 'reported very close correlation and no difficulty in coming to an agreed mark'.

This result is to be compared with that of Hitchman who reported a mean r, for Prose Reading and Conversation combined, of 0·78.

The English Speaking Board included in its test-battery (1) Prepared Reading ('prepared' in the sense that the candidates had read *and liked* a book from which the examiner chose a passage), and (2) 'Own choice' prose reading (optional to poetry and drama). Since marks for these test items were added to those for the other test items in the Board's battery before correlations were computed, it is not possible to show correlations comparable with those of Southampton or Hitchman.

(5) *Poetry Speaking*

The only experimenter so far to deal with Poetry Speaking is Hitchman who found a mean reliability coefficient r (for colleges only) of 0·79.

(6) *The Talk*

This last item was used by Southampton and the English Speaking Board (in schools) and Hitchman (in colleges). Southampton reported that the *Prepared Talk* correlated most highly (i.e. by comparison with the other types of test) with teachers' assessments, and 'on this evidence together with the enthusiasm shown for it by the external examiners, it may be reckoned the most successful'. Validity coefficients were approximately 0·55 and reliability coefficients ranged from 0·837 to 0·59. Hitchman found a mean r of 0·77 (reliability).

The English Speaking Board considered that the Talk (called by the Board 'Personal Project') was the most important and meaningful section of the examination. It certainly had several points in its favour: candidates could deal with any subject in which they had a keen interest; they could use visual aids carefully organized beforehand, employ a friend if they wished as a non-examinable collaborator, and have the benefit of full-scale audience reaction and participation. However, if the aim is to test the candidate's ability to handle spoken language, the test, as used by the English Speaking Board, has one grave disadvantage—the talk can be learned by heart.

(7) *Number of Assessors*

The number of assessors and their exact functions varied considerably in the four research studies under review. Using tape-recordings for his Listening Comprehension research, Wilkinson could use as many assessors as he wished. The Southampton experimenters used for each test method one examiner-assessor, one observer-assessor, and one tape-recording assessor (except for the Group Discussion method in which tapes were not used). Hitchman used in his Type A Tests (in which candidates were tested once only) one non-assessing examiner and two, three or four observer-assessors, and in each of his Type B tests two examiner-assessors (each candidate being examined twice with 'equal and parallel' test material, once by one examiner in one room and once by a second examiner in another room). Each of the English Speaking Board experimental tests was conducted by one examiner-assessor, who had with him or her an invited group of 'moderators', in number from four to ten. For each candidate the average of the moderator's marks (out of 25) was added to the examiner's mark to give a final mark out of 50. (The correlation coefficients of the examiner's marking v. the moderators' average ranged from 0·96 to 0·56.)

Both Southampton and Hitchman agreed that it would be a good thing to have more than one assessor at any one test. Southampton asserts that 'a double assessment leading either to an agreed mark or an average mark would seem to be essential. If necessary the second assessment can be made after the examination of tape-recordings has been made'. Hitchman (1966a) prefers the average mark. 'If two or more assessors are employed it is essential to take the average of their several judgements, made *in isolation* (i.e. without discussion). If each candidate is discussed various unintended weightings may be embodied in the mark(s) awarded. The most influential assessor may, for example, distort the judgements of his colleagues.'

While a test authority might be able to afford (in terms of fees or teachers' time) two assessors for any one test it is unlikely that it could afford more than two. The English Speaking Board claims that 'the close agreement in marking between examiners and moderators makes it reasonable to affirm that all are marking the same thing'. However, the Board reports poor agreement amongst moderators with no previous examining experience (as one might expect) and between *individual* moderators and the examiner in many cases. What the 'close agreement' does suggest is that several moderators, perhaps 4 to 10, are needed to get a reliable result for any one test.

(8) *Number of Gradings*

The English Speaking Board required its assessors to use a five-point scale 1-5, Southampton a ten-point 1-10, and Hitchman an eleven-point 0-10. The C.S.E. authorities use a six-point scale 1-5 with a sixth (called 'unclassified'). The American researcher Thompson (1943) devoted a large part of his study to the attempt to discover the best grading instrument for Speech-Making. He examined seven American grading systems and came to the conclusion that none was really superior to the rest. His study indicated that 'to improve rating speech educators must devote their time to the individual rater, and not to the method by which rating is done'.

Both Southampton and Hitchman noted a general failure on the part of their assessors (and the English Speaking Board a failure on the part of some) to use the full number of gradings available. This is a well-known state of affairs in all subjects where subjective ratings are used. Hitchman concluded that to get a true 5-point scale raters must be asked to use an 11-point 0-10, the resulting score then being arranged into five gradings. Southampton reported that tape-recording enabled their assessor to spread his marks more than the 'live' examiner and observer could.

(9) *General Impression Marking*

The Southampton committee, before their experimental testing took place, decided that the principle of rapid subjective marking often employed in written work* could and should be extended to oral examinations. They did, however, attempt to balance general impression with analytical marking by asking their assessors to study lists of itemized 'qualities' and to bear them in mind when making their assessments. Hitchman's method was to ask his assessors first to judge a number of 'details' (or qualities) and then to give a final single General Impression mark (which he considered especially valuable since it had been steadily built up in the mind of each assessor as he had been rating the particular 'details').

---

* and championed by Wiseman (1949) in the marking of English Composition for grammar school selection.

The English Speaking Board examiners awarded a single general impression mark, the moderators (whose average mark was added to that of their particular examiner) a series of 'detail' or 'quality' marks for *voice*, *articulation/pronunciation*, *vocabulary and composition*, *audience contact and relationship*, and *imagination/sincerity/spontaneity*.

(10) *Fatigue Effects in Assessors*

Using two comparison techniques Hitchman examined the marks awarded by fifteen assessors in six days' examining in schools and six in colleges. There appeared to be no fatigue effects.

Conclusions

(i) The tests considered in this article are reasonably valid and reliable (at least as valid and reliable as the composition test of the written-essay type), although it would be better if test authorities could afford *two* assessors.
(ii) To some researchers and some test authorities the Talk appears to be the most attractive and most successful test-item.
(iii) It is perfectly feasible to include a test of *Listening Comprehension* in a battery of spoken English Tests.
(iv) The best method of marking appears to be by *General Impression* combined in some way with *analytical* marking.
(v) Assessment of tape-recorded test material appears to be a practical proposition.

The future

It is obvious that we need to know a great deal more about tests and testing, and not only in the areas of validity and reliability. We must experiment with other forms of oral English testing, including those associated with poetry speaking and drama. Among the questions to which we must get answers are the following:-

(i) What battery of tests gives the most useful and accurate information about pupils' oral English?
(ii) What is the place of the tape-recorder in the assessment of oral English?
(iii) What is the best method of scoring?
(iv) How many categories ought a competent assessor to be asked to discriminate?
(v) Are there any fatigue effects in assessors; if so, how can their operation be modified or eliminated?

(vi) Can the judging of assessors be improved with training?*

Written and spoken English, we know, permeates the curriculum. Sooner or later we shall have to consider how to test *orally* such regular school subjects as mathematics, physics-with-chemistry, biology, geography, history and perhaps others.

References

Henry, E. (1964). *Widening Horizons: a report on the exploratory examinations in Secondary Schools carried out by the English Speaking Board.* Liverpool: The English Speaking Board.

Hitchman, P. J. (1963). 'Studies in Speech Measurement: An Experimental enquiry into the Validity and Reliability of Tests in Spoken English in Schools and Training Colleges'. Unpublished thesis, University of Nottingham.

Hitchman, P. J. (1964). 'The testing of spoken English: a review of research'. *Educational Research*, VII, I, 55-72.

Hitchman, P. J. (1966a). *Examining Oral English in Schools.* London: Methuen.

Hitchman, P. J. (1966b). 'The validity and reliability of tests of spoken English'. *British Jnl. of Educ. Psych.*, 36, 15-23.

Schools Council. (1966). *Certificate of Secondary Education: Trial Examinations—Oral English.* (Examinations Bulletin No. 11). London: H.M.S.O.

Thompson, W. (1943). 'An Experimental Study of the Accuracy of Typical Speech Rating Techniques'. Unpublished thesis, North Western University, Evanston, Illinois.

Wilkinson, A., Davies, A. *and* Atkinson, D. (1965). *Spoken English.* Birmingham: *Educational Review*, Occasional Publications, Number Two.

Wiseman, S. (1949). 'The marking of English composition in grammar school selection'. *British Jnl. of Educ. Psych.*, 19, 200-209.

---

*Thompson (1943) uses the word 'improve' to mean inducing a movement in assessors' marking towards the *mean*.

RESEARCH NEWS*

Edinburgh University. *The New Godfrey Thomson Unit for Educational Research*

During the years when the late Sir Godfrey Thomson held the Chair of Education at Edinburgh University, he developed and enormously expanded the research into psychometrics which had already earned him fame as a pioneer.

Before his death in 1955, Sir Godfrey established a Research Fund, which he saw as a means of perpetuating both the production of new psychometric tests and the sort of active inquiry which demanded the support of statistical as well as educational expertise. During the next ten years additional members of staff were appointed as Research Fellows and Assistants, much exploration into test content and methods of appraisal was undertaken, and the advisory service to Local Authorities using the Moray House tests developed markedly, often taking the form of an experiment or investigation done in collaboration with the Authority concerned.

In November of last year the University decided to give this part of the work of its Education Department more explicit recognition and better facilities by setting up, in separate premises, the Godfrey Thomson Unit for Educational Research. This Unit now housed at 10-12 Buccleuch Place, Edinburgh 8, is being directed by Dr. A. E. G. Pilliner.

How those responsible for it see its immediate future may be briefly stated in three parts. Firstly, the construction of tests will continue, though it is expected that, as selection procedures are modified or abolished, more time will be devoted to experiment with new types of test, and to more painstaking investigations of their usefulness. Secondly, the staff hope to continue and expand their collaboration with Local Authorities. Thirdly and it is perhaps in this respect that the establishment of the Unit as a separate body will have most meaning—it is hoped that not only other Departments in the University but also outside bodies or individual workers who want to do research in the social sciences and especially in the educational field will find that the resources of the Unit can offer many kinds of valuable help.

Educational research centre. *St. Patrick's College, Dublin*

St. Patrick's College in Dublin has set up an Educational Research Centre with the support of the Department of Education of the Irish government. The Centre, which is the first of its kind in Ireland, began operation in January 1966. It is staffed by two psychologists: Dr. Thomas Kellaghan and Mr. Liam Gorman.

The Centre plans to carry out fundamental research in child development, paying particular attention to the development of cognition. It will also study problems peculiar to education in Ireland. Some work will be done on the construction and standardization of tests suitable for use in Irish schools. While in the past, there has

---

*The editors are pleased to receive information on the development and progress of research centres and projects, which is likely to be of particular interest to readers of *Educational Research*.

been a number of isolated empirical studies of child psychology and education in Ireland, the establishment of this Centre should both widen the scope of, and make for greater continuity in, research efforts in these areas.

The fact that the Centre is located in a college that is concerned with the education of teachers should ensure that there is close contact between it and the teaching profession from the start. It is indeed hoped to interest and involve teachers in all of the work undertaken by the Centre. Their co-operation will be sought in the planning and execution of research, and they will be provided with information on findings. The inclusion of teachers is seen as having two effects. Firstly, it will increase considerably manpower resources for research; and secondly, it will ensure that the people most likely to be able to act on research findings—the teachers themselves—will be aware of, and feel involved in, the research.

It is proposed that the College should launch an educational journal, possibly during the coming year, and that issues would appear twice a year. In this event, the Research Centre would utilize this medium for the publication of reports on the progress of its research work. The Research Centre would also issue individual *ad hoc* reports, from time to time.

A new journal. *Sociology of Education Abstracts*

It has become apparent during recent years that educationists must, if they are to be effective, take greater account of sociological research. Reports and articles of a sociological nature which are particularly relevant to the study of education, have been produced in increasing numbers; but these have been until now largely lost to the educational world since an effective abstracting service has been lacking.

Now, thirty-five specialists in education and sociology have organized themselves into a reviewing team which will abstract articles from 150 major social science and educational journals throughout the world. The abstracts will be published in a quarterly journal, *Sociology of Education Abstracts*: and the first issue of Volume I, which contains 117 abstracts of articles published during 1965, is now available.

Each abstract gives a full bibliographical reference, together with a summary of 200/300 words which will enable the reader to identify areas of interest and follow them to the original if required. Four quarterly issues make up a volume, and each volume will have indexes for both sociology and education.

Sociology of Education Abstracts is available to institutions at 2 guineas ($7) per volume, and to individuals at 1 guinea ($3). Copies of the first issue may be obtained on approval from: The Business Manager, Sociology of Education Abstracts, Department of Adult Education, University of Liverpool, 9 Abercromby Square, Liverpool 7.

This publication fills an obvious need, and is to be commended to libraries, to research workers, and to all educationists and social scientists who wish to keep abreast of developments in the sociology of education.

Research Note:

THE EFFECTS OF EARLY ENVIRONMENTAL STIMULATION AMONG FREETOWN CREOLES—A COMPARATIVE STUDY

by O. M. Ferron

Department of Education, Ahmadu Bello University

Introduction

STUDIES like those of McGraw in Florida in 1931 have suggested that there are fundamental differences in motor and intellectual development between infants and young children of different races, but such results have been refuted by subsequent investigators like Brown (1944), Gilliland (1951), Anastasi and D'Angelo (1952) and Pasamanick and Knobloch (1953). If anything, these later studies have pointed to a slight superiority in motor development in favour of children of African descent. Thus in 1956 Akim *et al* found that the majority of African children whom they tested were in advance of European norms in motor development at similar ages. In an investigation in Zaria, N. Nigeria, in 1963, 80% of the respondents to a questionnaire on child rearing practices agreed that ' most African children learn to walk at the age of one year or less ', while 10% were undecided. According to some authorities the corresponding age for British children is from 12 to 14 months. The investigations of Geber and Dean would appear to require further scrutiny and confirmation, but, using the Gesell Schedules, they did in fact report African developmental quotients as high as 360! These authors concluded: ' The results of the testing showed an all round advance of development over European standards which was greater the younger the child'. The finding that the gap between the intellectual performance of children of European and African descent respectively widens with age, has been confirmed by several studies, and I.Q. differences of as much as 20-30 I.Q. points have been reported when secondary school children of different races in the U.S.A. and West Africa have been compared. The general consensus of opinion now is that there is little or no evidence for ethnic differences of a hereditary nature, and that most of the large differences observed can usually be accounted for in environmental terms. Moreover, most psychologists are now agreed about the fundamental importance of early environmental stimulation for subsequent intellectual development, so that it is a matter of some importance for educators in Africa to discover at exactly what stage in the child's development differences between African and European performance become significant.

The Freetown experiment

The subjects of the present study were a group of 30 Creole nursery school children in Freetown who were matched on the basis of age and socio-economic level with a group of 30 English children in a nursery school in Deptford, London. A purified pre-school test consisting of 25 items categorized into the following groups was drawn up after some preliminary experimentation and statistical procedures:

Group I: Perception and Muscular Co-ordination

			Discriminating Value†
1.	Item No. 1*	Formboard	(.40)
2.	,, No. 6*	Building a tower	(.70)
3.	,, No. 7*	Nest of beakers	(.75)
4.	,, No. 8*	Montessori's pink tower	(.85)
5.	,, No. 17*	Montessori's cylinder block	(.45)
6.	,, No. 18*	Montessori's long stair	(.40)
7.	,, No. 19	Building a longer tower	(.80)

Group II: Perception

1.	Item No. 12*	Cut halves of a picture	(.60)
2.	,, No. 14	Matching geometrical figures	(.25)
3.	,, No. 15	Colour matching	(.35)
4.	,, No. 20	Matching solids	(.10)
5.	,, No. 21*	Threading beads	(.45)
6.	,, No. 22	Which jar contains more water	(.35)
7.	,, No. 23*	Matching geometrical figures	(.70)

Group III: Muscular Co-ordination

1.	Item No. 4	Cutting with scissors	(.65)
2.	,, No. 5	Drawing up a string	(.50)
3.	,, No. 13	Folding paper	(.25)

Group IV: Drawing and Writing Skills

1.	Item No. 2*	Garden path	(.80)
2.	,, No. 3*	Copying a circle	(.50)
3.	,, No. 16	Copying a cross	(.65)
4.	,, No. 24	Copying a square	(.85)
5.	,, No. 25*	Copying a star	(.55)

Group V consisted of three language items (Nos. 9*, 10, 11*) in which the children were required to identify, categorize and name miniature objects spread out at random on a table before them. Most of the children in Freetown were tested on an individual basis by the investigator himself, but the children in London were given the tests by a small group of students from the Rachel McMillan Training College in Deptford under the author's personal supervision. Every precaution was taken to ensure a standardized procedure in administering the tests.

---

† Upper and Lower thirds

*In the case of these items, differences between European and African performance were statistically significant.

The figures above in brackets indicate discriminating values, where the maximum value possible is 1. The split half reliability coefficient of the test was found to be .85 which when corrected by the Spearman-Brown formula gave a value of .92. The validity of the test can be justified partly by the fact that it was made up of items modified from the tests of well known authors. Moreover, according to Anastasi and Foley: '..... The degree to which ..... test discriminates between successive age groups, as well as the correlation between chronological age and score, are taken as indices of validity'. This is in fact a major criterion for the validity of such pre-school tests. With one or two exceptions, the test under consideration seemed to differentiate fairly well between younger and older children. The correlation between test score and chronological age (N=70) was found to be .71.

Results

The following results were obtained:

1. The mean score of the British children was found to be 19.4, that of the Africans 15.2, out of a possible score of 25. The difference of 4.2 points was statistically significant at the P=.01 level.
2. The children were divided into groups on the basis of age: 3 - 3:5, 3:6 - 3:11, 4 - 4:5, 4:6 - 5. At every age level the mean score of the British children was higher than that of the Africans.
3. The performance of the British children on all five sub-tests was superior to that of the Africans, the level of significance being P=.001, .05, .001, .05 and .05 respectively.
4. Significant differences in European and African performance were obtained in respect of the following items:
   Nos. 1, 2, 3, 6, 7, 8, 9, 11, 12, 17, 18, 21, 23 and 25.
5. In the following items the differences were not significant:
   Nos. 4, 5, 10, 13, 14, 15, 16, 19, 20, 22 and 24.
6. The Africans obtained a higher score only on item No. 5—' Drawing up a string ', but this difference was not significant.

Discussion

Of the significant items Nos. 1 and 23 relate to the perception of geometrical shapes. Ten of the younger African children failed on item No. 1, while only one European child did so. In the case of item No. 14 the difference is not significant; it becomes significant again in item No. 23 in which more unfamiliar shapes are involved. This would suggest that the observed differences might possibly be due to the fact that the young African child has less experience of even simple shapes like the square, circle and triangle than his European counterpart. As he grows older deficiencies may be made good, but the richness and variety in a western environment probably again give the European child an advantage with increasing age.

Item Nos. 2, 3 and 25 are concerned with drawing and writing skills, and it was found that the largest difference occurred in item No. 2—‘Garden path’. This may be partly explained by the fact that pre-school African children do not have as much practice as British children in handling a pencil, since they are not normally encouraged to write and draw at home. Failure in the ‘Garden path’ was invariably due to lack of control in keeping the point of the pencil along a certain path, so that obviously children who have had much practice in scribbling, and who proceed from this random activity to more controlled ones such as shading in printed outline drawings, etc. would develop a higher degree of skill than those who are denied these activities in the home.

The remaining significant items, with the exception of Nos. 9, 11, 12 and 21, relate to a rather fine discrimination of dimension. It should not be difficult to appreciate that those children who are showered with toys of all shapes and sizes practically from birth, would have an advantage in this respect. In the last few years a particular kind of toy of which the ‘Nest of beakers’ in item No. 7 is typical, has become very popular in British homes. Barrels, cups, tea-pots, Easter eggs, etc. become progressively larger, so that this type of toy not only caters for the child's natural tendency for putting things inside each other, but it also gives a good deal of practice in the fine discrimination of dimension. Item No. 12 relates to pictorial representation in which other investigators have also found children in non-western civilizations to be deficient, while in item No. 21, obviously the provision of toys from an early age would give the British child an advantage.

The relatively small differences in the case of the language items may have been partly due to the fact that the ceiling was too low for the British children. On the other hand, the general impression gained was that while the African children lacked the coherent conversational ability and ‘chattiness’ of British children, they did not strike the investigator as being deficient in vocabulary. This is confirmed by Carroll's West African survey in 1962 in which he found that in the English vocabulary tests he administered, African children obtained scores ‘very little lower than those of English children’, and their combined English and vernacular vocabularies were rather higher than English norms. In fact the linguistic abilities of Africans generally can hardly be gainsaid. Thus through force of circumstances perhaps, many West African children can speak three or four different languages, and there are grounds for believing that the observed deficiencies in the use of the English language are to a large extent due to poor teaching.

Another observation seems worthy of mention and that is that, although the Creole group appear to be ‘inferior’ to the British group in ‘Drawing and Writing Skills’, item No. 16—‘Copying a cross’ showed the smallest difference (not significant) of all the items in this category. A possible explanation is that the Creole culture is based on Christianity, and practically every Creole village has its own chapel. The village preacher and the village schoolmaster appear to be as important influences in the lives of young Creoles today as they were among the village communities of 18th and 19th century England. It is likely therefore that the cross is one of the shapes the Creole child perceives very early in life, and it may be that it has an important emotional significance for him.

Conclusions

The present day Creoles of Sierra Leone are the descendants of former slaves from the U.S.A., Canada and Britain, and those who were re-captured on the high seas after the abolition of slavery by the British Parliament. These homeless Negroes were settled in the area that eventually became known as the 'Colony', now re-named the Western Area of Sierra Leone. In ancestry and culture they would thus appear to have much in common with Negroes in the U.S.A. It was therefore rather surprising that while similar studies in the States showed no statistical difference between the performance of 'white' and Negro children at the pre-school level, the differences between Creole and British children should be of the order specified. Now 'Western' peoples did not originally set out to provide toys for their children in order to aid their intellectual development. They did so primarily because it was a recognized and culturally acceptable way of displaying their love and affection for their offspring. On the other hand, it is a well known fact that in certain African and Asian cultures rather more emphasis is placed on the sensual component of parental love, to the comparative neglect of the symbolic aspect involved in giving and receiving. In this connection the American Negroes and the Sierra Leoneans differ in one main essential: while the former have assimilated 'white' culture to a greater degree, the latter have imitated it rather more mechanically because of the peculiar circumstances in which they found themselves. The social pressures brought to bear on the American Negro are essentially pressures from a 'white' community so that, for the most part, they do what 'white' people do, and their values are typically 'western'. The social pressures brought to bear on the Creoles are pressures from a 'black' community, and their values tend to be rather ambivalent.

In investigations in Sierra Leone and Nigeria from 1960-64, the present author was unable to find any conclusive evidence regarding ethnic differences of a hereditary nature, thus confirming the results of previous research in this field. Hence the observed gap in intellectual performance between Africans and Europeans at the pre-school, primary and secondary school levels is probably almost entirely due to the nature of the environmental stimulation provided. Much has been and is being done at the higher levels, but relatively little at the pre-school and primary level. While a concerted attack must be made on many fronts, might it not be possible that better results could be expected if much greater attention than hitherto were given to the laying of sound foundations by, say, the large scale provision of good nursery schools to counteract the effects of poor home environment, and the provision of well trained teachers, especially for the lower classes in primary schools? Investigation in Freetown revealed the fact that the youngest children in primary schools seem to be allotted the worst classroom accommodation, the most dilapidated school furniture ill adapted to their size, the minimum of apparatus, and almost certainly the least qualified teachers!

References

ANASTASI, A. and FOLEY, J. P. (1949). *Differential Psychology*. Rev. ed. New York: Macmillan.

ANASTASI, A. (1958). *Differential Psychology*. 3rd ed. New York: Macmillan.

CARROLL, B. J. (1962). 'An English language survey in West Africa', *English Language Teaching*, Vol. XVI, No. 4.

FERRON, O. M. (1964). 'A Study of Certain Factors Relating to the Tested Intelligence of Groups of Some West African Children'. Ph.D. Thesis, University of London.

GEBER, M. and DEAN, R. S. A. (1957). 'Gesell tests on African children', *Pediatrics*, Vol. 20, No. 6, 1055-65.

GEBER, M. (1958). 'The psycho-motor development of African children in the first year, and the influence of maternal behaviour', *Journal of Social Psychology*, No. 47, 185-195.

McFIE, J. (1961). 'The effect of education on African performance on a group of intellectual tests', *Brit. J. Educ. Psychol.*, 31, II, 233.

Research Note:

SCHOOL REFUSAL: AN INQUIRY INTO THE PART PLAYED BY SCHOOL AND HOME

by MARGERY G. COOPER

Institute of Education, University of Durham

A SUDDEN increase of interest in the problem of school phobia was accentuated by letters and articles in the press. In November 1960 the *Observer* wrote: 'School Phobia is on the increase and some of the statistics make it look a formidable problem'. An inquiry therefore was set in motion. A letter accompanied by a table outline was circulated among the Child Guidance Clinics and School Psychological Services in England to all Directors of Clinics, explaining the nature and purpose of the inquiry. The table outline asked for information on the number of children referred for (*a*) school phobia; (*b*) truancy, in relation to:

1. The total number of cases referred to the clinics in each of 5 years, 1956-60.
2. The numbers of school population of the clinic area in those years.

101/178 Clinic Directors replied expressing interest, but only 38 were able to give information. Others gave reasons for abstaining, most important of which were pressure of work, lack of a defined catchment area, and a reluctance to distinguish between cases. Of the 38 completing the table, only one expressed complete confidence in the information given. Very few figures for truancy were supplied. The main reason given was that truants were referred to School Welfare Departments where records were not kept.

Information received showed that school phobia as a % of all clinic referrals had increased from 4.1% to 6.2%, and that the number per 10,000 of the school population had grown from 1.36 to 2.76 over the five-year period. The upward trend, as Eisenberg (1958) remarked, 'probably reflected a change in the recognition and referral from physician and school authorities'.

The following comments from Directors illustrate the complexity of the problem and the difficulties inherent in compiling data.

'We have very few cases of truancy referred as this is often a social problem and is dealt with by the School Welfare Officer'. Diagnostic difficulties are well represented: 'The single diagnostic symptom of truancy is rarely made—the diagnosis always takes into account various allied conditions', or 'it is impossible to give the numbers of cases of truancy as it is very rare for children to be referred for this symptom alone whereas it does occur, particularly in those cases referred from the courts'. There was much more comment on the incidence of school phobia: 'a psychiatric fad', 'I find that school phobia is an increasing problem and I feel that teachers are mainly responsible for creating it'; but again many comments revealed reluctance to classify: 'We do not classify cases and we would not classify by symptoms', or 'We have considerable reservations as to the usefulness of the division into two opposing categories which overlap and extend into other syndromes', or 'We do not code School Phobia as a separate item'. Finally, in spite of, 'There are truancy cases which are better regarded as school phobics', 'This exhausting yet interesting problem . . .', the research developed, perhaps because of the comment, 'You are on to a problem about which very little is known'.

The aim of the investigation reported in this paper was to inquire into the part played by school and home in the problem of school refusal—a comprehensive term designed to include all children persistently refusing, for whatever reason, to attend school.

It seemed from previous research that such children fell broadly into two groups:

(a) those where refusal to attend school was thought to be a manifestation of a wider psycho-neurotic syndrome associated with the parent/child relationship, having been referred to the School Psychological Service as school-phobics.

(*b*) those, where refusal took the form of persistent absence without lawful reason, who were not attending the School Psychological Service but were the concern of the School Welfare Department as truants.

It was recognized that some cases might not fall clearly into either group but previous research appeared only to stem either from Psychological Services or Welfare Departments. There was no information on school life or how these children appeared to their teachers. In many cases knowledge of diagnosis and treatment was withheld from school personnel. It was decided therefore, (*a*) to make some investigation into how far the school could be said to play a significant part in this refusal; (*b*) to examine some assessments made by the schools of certain personal abilities and characteristics of these children as they appeared to the teachers compared with those of children attending regularly; (*c*) to estimate how far different aspects of school refusal were related to the home, as distinct from the school life, of these children.

Procedure

Two groups of 40 children each were formed, Group A being cases from the files of a School Psychological Service and Group B from School Welfare. A condition of selection was that the children should still be of school age. Visits were paid to the schools and two control groups AC and BC—children of the same age and sex who were first on the register of the respective classes—were also established.

Because of movement of children between schools, a total of 97 schools were visited. Prepared interview schedules were drawn up to elucidate information about the school from the Headteachers, the child in school from the teachers, and the child's home as known by the school from school personnel. Care was taken to render the interview technique valid through the adoption of schedules, the establishment of rapport and the control of contextual effect. All the assessments were subjective but this was not felt to lessen their importance.

The schools

Points of inquiry centred on size of school and class, methods of organization and teaching, the use of rewards and punishments and standards of home and school relationship.

Table I

	Size of School			*Size of Class*		
	0-100	101-300	300 +	0-30	31-40	41 +
A	—	35%	65%	32½%	57½%	10%
B	—	10%	90%	50%	47½%	3%

Both streamed and unstreamed classes had cases of refusal. 37·5% A, 33% B were in unstreamed classes and the percentage of unstreamed classes in the groups was higher than the over-all percentage in the total area investigated. 25% A, 12½% B were in classes conducted on informal lines: 25% A, 33% B were in schools where corporal punishment was administered. The use of tangible rewards, awarding of tokens, operated in a high percentage of schools for both groups, and standards of home and school co-operation were judged average or good.

The attitude of the schools towards Group A revealed a wish to carry out adjustments to meet expressed fears, yet only 60% of Headteachers were ready to change a child's class, or give permission not to attend certain classes, feeling that the framework of school life had been designed for the support of even those children finding difficulty in attending. Among adjustments found in schools however, was one boy permanently excused Physical Education, another where special arrangements were made for eating the school meal, a third had moved into another class and a fourth did his work in the Headmaster's study. Class teachers who had been involved in the more serious cases spoke of their personal distress in seeing children break down so completely. On the matter of return to school, Headmasters felt this should be unobtrusive and that the onus of seeking remedial help to offset time lost in school should be placed on the home.

Attitudes towards Group B were notably less sympathetic. The close co-operation between school and Welfare Officer had reduced truancy to spasmodic inattendance, but the children were regarded as ' offenders '. 10% were persistent in their refusal and had background situations affected by neurotic parents, thus presenting a home situation comparable to Group A, but these had been passed on by the Psychological Services and had become the concern of the Welfare Officers. The complexity of the cases made prognosis doubtful.

The type of school at referral showed a high number of Grammar School children in

Group A and this group, further investigated, gave additional evidence of breakdown at times of change of school or class.

Table II: Type of School at Referral

Type	*Group A*		*Group B*
	Referral	*Interview*	*Interview*
Infant	6	2	2
Junior	13	8	11
Primary	5	3	1
Secondary Modern	7	15	23
High	1	2	2
Grammar	8	8	1
Special	–	2	–
	40	40	40

Mean Age G. 8.8 N=19   12.7   12.5 N=19
B. 9.8 N=21   12.2   12.5 N=21

Table III: Concomitance of Breakdown with that of a change of School

	Group A	
	Boys	*Girls*
At transfer		
5 years Infant	1	5
7 ,, Junior	2	1
11 ,, Modern	4	1
11 ,, Grammar	2	3
14 ,, Grammar	–	1
Pre-transfer		
7 years Junior	–	1
11 ,, Junior	4	–
14 ,, Modern	1	1
	14	13

15% could not accept school at 5, 35% at time of transfers, 7 to Junior School, 11 from Junior to Secondary, and 14 Secondary to Grammar: 17½% pre-transfer—included one child at 7 afraid of the ‘ big ’ school, four at 11 preparing for entry to Secondary School, and two preparing for entry to Grammar (from Secondary). Previous investigators had stressed the precipitating factor of change of school or class. Furthermore it was established that of all groups A had endured more changes of school than all other groups.

Table IV:

Changes of School—in excess of number required within the educational system i.e. one change Infant-Junior and one change Junior-Secondary

	Infant		*Junior*						*Secondary*							*Total excesses*	*No excess*	*With excess*
	0	1	0	1	2	3	4	5	0	1	2	3	4	5	6			
Group A	1	3	9	3	2	–	–	–	9	9	2	–	1	–	1	33	19	21
Group AC	2	2	8	4	–	1	–	1	13	7	1	–	1	–	–	27	23	17
Group B	2	–	6	4	–	–	–	1	19	4	1	–	2	–	1	29	27	13
Group BC	2	–	9	2	–	–	–	–	15	12	–	–	–	–	–	14	26	14

Summary

There were no discernible differences between the schools attended by the two groups of children. Cases were found in schools where what might be considered progressive methods were practised. The differing attitudes of the schools towards the groups was apparent, there being more sympathy expressed towards the children with a phobic reaction. There was evidence to corroborate previous findings that refusal to attend school by such children took place at times of change of school or class.

The abilities and personal characteristics of the children as they appeared in their schools

Data available in the schools concerning the children's intelligence attainment covered a wide field and standardization for attainment was possible only within a group N=17.

Table V: (N.F.E.R. Group Tests)

Group	*Mean I.Q.*	*Range*	*Mean Attainment Score*	*Range*
A	104.1	76–122	105.7	80–133
AC	105.3	84–126	103.3	85–122
B	85.3	65–103	87.2	70–107
BC	84.4	70–100	85.5	60–106

The evidence added weight to previous findings that intellectual or educational retardation does not appear to be a relevant factor among children with a phobic reaction to school. Truants, in a lower intelligence range were here found working up to standard. The work of the Welfare Officers in reducing periods of truancy could have affected this lack of retardation.

An assessment of physique revealed Groups A and B less well developed than their fellow pupils (P =.05), while in appearance Group B stood out from all other groups as being uncared for (P =.01).

Behaviour traits of timidity, lack of persistence, sociability, fastidiousness, attitude to failure, success, and truth, were marked on a scale by the teachers and in order to test further the social reactions assessed by the teachers inquiries of the children themselves were made of attitudes to school, membership of clubs and out-of-school activities.

Groups A and B . . .
Group A differed significantly (P =.01) from B in the following ways:

They were—more intelligent
more truthful
more anxious to respond to authority
more persistent
more fastidious
more affected by failure
They clung more to home in out-of-school activities.

Group A did *not* differ significantly from B in the following ways:

poor physique
timidity
anxiety to be useful
fantasy
inability to enjoy success
withdrawal

Groups A and B had less contact
(*a*) with the teacher
(*b*) with their peers than the other children and did not belong to outside clubs.
They appeared not to enjoy school.

Summary

Patterns of behaviour in the control groups were remarkably alike and helped to highlight the similarities and differences between Groups A and B, who presented a common basis of partial withdrawal from social interaction. Their overt behaviour patterns included shrinking, timidity, seeking refuge in unreality-fantasy, remaining quiet, withdrawal from class activities and experience of success, and hiding, seeking little contact with teachers and peers. Thus they were ' different ' from other children. They appeared to employ a major defence mechanism of retreat. Subsidiary traits for the phobics, fastidiousness and reliance on authority, for the truants, untidiness and defiance of authority, seemed to be related to conditions prevailing outside of school.

The home life of the children

Assessment of the home life by the school was supplemented by information gained in the interview with the children. Allowance was made for the unreliability of the children's estimates and new information gained in this way was discussed with the Headteacher. Points of inquiry were the fathers' occupations, home district, size and position in family, parents' attitudes to home discipline and support and encouragement in school.

Group A differed significantly ($P = .01$) from B in the following ways:

The parents occupied a higher socio-economic level.
Families were smaller.
Home discipline was over-anxious.

Attitudes to education and support at school were better. Group A were distinguished ($P = .01$) from all other children by 'over-anxious' discipline administered by a dominant mother. The phobics and truants appeared to occupy the opposite extreme position on the range of socio-economic status and the differences in size of family and attitudes to education probably stemmed from this.

Conclusion

The role of the school in School Refusal appeared to be a minimal one. Children refusing to go to school tended to be socially withdrawn, employing defence mechanisms in order to establish themselves in a situation which they did not particularly enjoy. Their mode of refusal, an anxiety reaction in the case of the phobics, and deliberate unlawful absence in the case of the truants, was seen to be related to their home background. The higher social level of the homes of the phobics appeared to help the children to absorb school values, but their anxiety, related to a dominant mother in an over-anxious home, irrupted at times of stress. The truants frequently exhibited a pattern allied to their social level. With a background of unconcern they tended to cease to identify themselves with school, being defiant and careless of its efforts. They stayed away, manifesting their need of security by a more positive, deliberate act of defiance. Both groups of children gave evidence of their need for specialized help.

The findings seem to suggest that there is much emotional ground common to both types of 'school refusal' reaction.

Research Note:

AN INVESTIGATION INTO THE USE OF THE S.R.A. READING LABORATORY IN THREE MIDLOTHIAN SCHOOLS

by H. B. Pont*
Department of Psyshology, University of Edinburgh

Introduction

THE Science Research Associates Reading Laboratory (II c)† is simply a large box containing reading material printed on separate sheets of cardboard. It introduces pupils to 'power builders', 'rate builders', and 'listening skill builders'. A 'power builder ', designed to increase comprehension in reading, includes an article to be read and questions on that article under the headings, ' How well did you read? ' and ' Learn about Words '. A ' rate builder ', to increase the speed of reading and improve concentration, consists of a short passage to be read in a limited time and a comprehension check. With a ' listening skill builder ' the teacher reads a passage, the pupils listen and complete a comprehension check, and the teacher reads out the answers. There are 144 ' power builders ' twelve at each of twelve levels, each level being designated by a different colour; 144 ' rate builders ' with the same distribution; and 10 ' listening skill builders '. Each pupil receives a printed record booklet in which he records his answers and results.

The reading laboratory is designed to be used in the normal classroom situation with a whole class, each pupil working on his own. No equipment, other than the material in the box and the record booklets, is necessary. It is largely self administered. The starting level for each pupil is determined from a starting level guide and thereafter group leaders give out material or pupils collect their own. Progress to a higher level occurs when the required level of success at the previous level is achieved.

In so far as the S.R.A. laboratory does not replace the teacher, though there is a certain amount of self-organization and self adjustment on the pupil's part, it is not a true teaching machine. It does however incorporate many features of a teaching machine, such as the need for an attentive, active pupil constructing his own response, the allowance for individual differences in the rate of learning, and the importance of knowledge of results.

This is the first such laboratory to be used generally in schools. Novelty is no criterion of worth, however, and it is necessary that developments of this type should be subjected to rigorous experimental assessment.

Design of the experiment

An experiment to find the effect of using the S.R.A. Reading Laboratory on reading quotient and verbal reasoning quotient was carried out in three Midlothian schools,

---

*Formerly Assistant Educational Psychologist, Midlothian Education Authority.
†Distributor: Thomas Nelson and Sons Ltd., London and Edinburgh.

Wallyford Primary School, King's Park Primary School, and Bryans Primary School. In all, two hundred and five children took part in the experiment; there were forty, in one class, in Wallyford, one hundred and twenty-five, in four classes, in King's Park, and forty, in one class, in Bryans.

In Wallyford the forty pupils, from primary 6A, were divided into two groups, a control group and an experimental group. These were matched for intelligence, the average V.R.Q. of the control group being 107 (S.D. =5.0 points) and that of the experimental group being 109 (S.D. =6.0 points); and for age, the average age of the control group being 10 years 8 months (S.D. =4 months) and that of the experimental group being 10 years 7 months (S.D. =5 months). Both groups were given a silent reading test (Schonell R4) before the experiment was begun, and from the results of this test a reading quotient, used to overcome the age change during the experiment, was obtained for each pupil. During the next twelve weeks the experimental group had their normal reading lesson replaced by the S.R.A. Reading Laboratory (II C) while the control group had their normal reading lesson. After twelve weeks the two groups were again given a silent reading test (Schonell R4) and a verbal reasoning test (Moray House Verbal Reasoning Test 58), and new reading quotients and verbal reasoning quotients were obtained for each pupil.

In King's Park Primary School four classes, primaries 4x, 4y, 5x, and 5y were involved in the experiment. Classes 4x and 5y acted as control groups, that is they had normal reading lessons, while classes 4y and 5x had the programmed reading instruction of the S.R.A. laboratory. In all, sixty pupils were in the experimental group and sixty-five in the control group. The groups were matched for intelligence, the mean verbal reasoning quotient of the control group being 105 (S.D. =14) while that of the experimental group was 104 (S.D. =15). All pupils were given a reading test (Schonell R4) before and after the use of the laboratory giving two reading quotients for each pupil.

At Bryans Primary School the reading laboratory was used with a class whose mean V.R.Q. was 89 compared with the mean V.R.Qs of 109 and 104 of the other experimental groups. No control group was used. All pupils were given a silent reading test (Schonell R4) before using the laboratory. Staffing difficulties caused the use of the laboratory to be discontinued in the eighth week, and work with it was not resumed till four weeks later. A second reading test was given when the full course of the reading laboratory was completed.

At the conclusion of the twelve-week experiment the groups returned to normal class reading. At Wallyford the two halves of the class were recombined and the reading book used was *New Worlds to Conquer*[1]. At King's Park School all classes involved used the appropriate books of the *New Reading* series[2]. At Bryans the reading books used were *Pleasure and Progress*[3], *Adventures of Reading*[4] and *Rainbow Readers*[5]. Thus the majority of the pupils were using reading books with a similar layout to the power builders of the S.R.A. laboratory, in that they include a passage to be read, direct questions on that passage, test comprehension, and extra work to do. The subject matter too was comparable being mainly factual in all cases.

---

[1] Ramsbottom, E. (1963). *New Worlds to Conquer.* Leeds: Arnold.
[2] Scott, A. F., ed. *New Reading.* London: Readers' Digest.
[3] Maxton, G. S., ed. (1958). *Pleasure and Progress.* Edinburgh: McDougall.
[4] Thomson, D., ed. *Adventures of Reading.* London: Black.
[5] *Rainbow Colour Books.* Glasgow: Collins.

About seven months after the completion of the experiment, all pupils involved in the experiment, both in the control groups and the S.R.A. groups, were given the reading test (Schonell R4) which had been given on the two previous occasions.

Results

(A) Results of 12-week Experiment

Table 1

Effect of S.R.A. Reading Laboratory on Reading Quotient

School	*Group*	*Mean Gain in Reading Quotient*	*Level of Significance of Gain*	*Difference between Groups*	*Level of Significance of Difference*
Wallyford	Control	8.8	.01	6.3	.05
	Experimental	15.1	.01		
King's Park	Control	8.05	.01	4.9	.01
	Experimental	12.95	.01		
Bryans	Experimental	9.5	.01		

Table 1 shows that all groups, experimental and control, made a significant increase in reading quotient over the period that the reading laboratory was in use. Although reading age would be expected to change with time, reading quotient, which contains a time factor, would not. This increase in the control groups may have resulted from the extra attention which was being paid to the pupils involved in the form of external examination and perhaps keener teaching. Alternatively it may have been due to the increased interest of the pupils, engendered by the knowledge that their reading was the focus of an experiment. However, in the schools where a control group was used it was found that the experimental groups made increases significantly greater than those of the experimental groups. However, it must be remembered that, while the attention of the experimental groups was being focussed almost entirely on silent reading, the control groups had their attention divided among several aspects of reading. Thus the control groups may have made significant increases in the other unmeasured aspects of reading though they fell behind in that field which was measured; namely silent reading for comprehension, as measured by Schonell R4.

Table 2

Effect of S.R.A. Reading Laboratory on V.R.Q.

School	*Group*	*Mean Increase in V.R.Q.*	*Significance of Increase*	*Difference between Groups*	*Significance of Difference*
Wallyford	Control	+2	NOT	1	NOT
	Experimental	+1	NOT		
Bryans	Experimental	+2	NOT		

There was no significant change in the mean V.R.Q.s of all groups for which data were available. As far as we are concerned in this experiment anyway, the Reading Laboratory does not have an effect on V.R.Q.

Table 3

Effect of V.R.Q. on the increase in Reading Quotient

Group	*Mean Increase in Reading Quotient*	*Significance of Increase*	*Difference between Groups*	*Significance of Difference*
V.R.Q. = 100 +	14·5	.01	3·6	NOT
V.R.Q. = 100–	10·9	.01		

In the analysis to find the effect of V.R.Q. on the increase in reading quotient, the pupils in the experimental groups were separated into two groups, those with V.R.Q. greater than or equal to 100 and those with V.R.Q. less than 100. This figure was taken as the dividing line as it was the lowest figure which gave two groups of size suitable for statistical analysis. It was found that the mean increase in reading quotient of the group with V.R.Q. greater than 100 was 3.6 points greater than that of the group with V.R.Q.

less than 100. This difference was not found to be significant however. This result does not exclude the possibility, however, that there may be a level of V.R.Q. below which the laboratory begins to have significantly less effect. That some such factor is at work is suggested by the results at Bryans Primary School. In Bryans, the experimental group had a mean V.R.Q. of 89 compared with those in the other experimental groups of 104 and 107. The mean increase in reading quotient at Bryans was significantly lower (at 0.05 level) than those in the experimental groups of the other two schools. This effect was complicated by the break in the middle of the experiment at Bryans, which might itself have resulted in a lower mean increase.

(*B*) *Results of Follow-up* (*7 Months after end of Experiment*)

Table 4

Change in Reading Quotient since the End of Experiment

School	*Group*	*Mean R.Q. at end of Experiment*	*Mean R.Q. at Follow-up*	*Mean Increase*	*Level of Significance*
Wallyford	Control	107.8	111.6	3.8	0.05
	Experimental	109.3	110.1	0.8	NOT
King's Park	Control	111.0	111.0	0	—
	Experimental	110.0	109.3	−1.7	0.05
Bryans	Experimental	98.1	98.7	0.6	NOT
King's Park	Exp.-Mean I.Q. = 110	119.0	118.9	−0.1	NOT
King's Park	Exp.-Mean I.Q. = 99	101.3	97.9	−3.4	0.05

Since the completion of the experiment the mean reading quotient of the control group has risen by 3.8 points while that of the experimental groups has not increased significantly.

In King's Park Primary there has been no significant change in the mean reading quotient of the control group but there has been a significant decrease in the mean reading quotient of the experimental group. Further analysis indicates that the decrease is made up of a decrease of 1.0 points in class 4y (a good class with a mean I.Q. of 110) which is not significant, and a significant decrease of 3.4 points in class 5x (a poorer class with a mean I.Q. of 99).

Table 5

Change in R.Q. since the Beginning of the Experiment

School	*Group*	*Mean Gain in R.Q.*	*Level of Significance*	*Difference between Groups*	*Level of Significance of Difference*
Wallyford	Control	12.5	.01	1.3	NOT
	Experimental	13.8	.01		
King's Park	Control	8.6	.01	2.6	NOT
	Experimental	11.2	.01		
Bryans	Experimental	9.2	.01		

The overall effect of the experiment in all schools has been to raise the mean reading quotients of all the groups involved. The mean reading quotients of the experimental groups were not found to have been raised significantly more than the control groups.

Conclusions

It is clear from the above results that the carrying out of this experiment has had a beneficial effect on the reading of the pupils concerned in the experiment. The mean reading quotients of the experimental groups and the control groups are higher than at the beginning of the experiment. However the results of the follow-up seem to suggest that this improvement does not derive specifically from the S.R.A. reading laboratory but more from the carrying out of an experiment.

While engaged on the experiment and immediately following it the S.R.A. groups did appear to be better readers. However in the months which followed the difference between experimental and control groups was almost eliminated. The methods by which this took place differed in Wallyford and King's Park schools. In Wallyford the experimental group remained static while the others gradually caught up, while in King's Park the control group improved less but the experimental groups actually lost some of their increased ability.

The reading laboratory is really a ' crash ' programme in silent reading instruction. It is a short twelve week programme and is not designed to be used continuously throughout the primary school. It boosts a child's reading ability. However every child has a ' ceiling ' of attainment at any particular developmental level, depending partly on personality and partly on intelligence, since obviously one cannot go on increasing the child's reading ability indefinitely. It would appear, then, that the reading laboratory simply enables pupils to reach their ceiling, or at least a higher level of attainment, more quickly and those not using the laboratory take, in the presence of an indirect stimulus, several months more. It is interesting to note that careful examination of the data collected indicated that, in general, pupils in all schools were working at a level below that which one would have predicted from their V.R.Q.s. This seems to suggest that with normal teaching methods, that is without the stimulus of an experiment or S.R.A. laboratory or similar apparatus children will work at a level below that of which they are capable. It is more easily seen, then, why the control groups were affected. Using this concept of a ' ceiling of attainment ', the King's Park results could be explained on the basis of over teaching. That is, the ability of the pupils has apparently, for a short time, been raised, perhaps by conditioning in a technique, above their actual ceiling. The lack of reason in the process then facilitated forgetting after the constant practice had ceased. This retrogression is similar to that which occurs, on occasion, with the poorer pupils of adjustment classes.

What then is the place of the reading laboratory in the primary school? It is not a substitute for the class reading lesson since it is too narrow. Concentrating on silent reading, it ignores the fields of recreative reading, oral reading and reading and interpretation. Nor does it appear to be better than the normal class lesson for improving silent reading, only quicker. However, the laboratory is, to a large extent, self-administered and puts the responsibility of organization on the pupils, giving the teacher the new role of ' learning consultant '. Any method of teaching which removes the burden of routine from the teacher is obviously worthwhile since it allows the teacher to discover and solve more individual problems. Also the scoring procedure is valuable in that it familiarizes the pupils with the detailed keeping of results and reading of charts. Its place in the normal class would then seem to be that of an extra whose main worth lies in the fact that it is ' different ' and thus arouses greater interest.

However, from what has been said above it can be seen that the laboratory is capable of raising the reading level quickly. If some child were reading well below his developmental level, that is if his reading age were well below his mental age, then the laboratory would be a very suitable method of instruction. Thus it would seem that the reading laboratory would be very suitable for use in adjustment classes where the aim is to bring a child quickly up to a level at which he can again join his class.

Research Note:

FIRST DEGREES

A Pilot Survey of Undergraduate Preferences Regarding Training, Assessment, and Choice of Career.

by A. R. Marsh

Peterborough Technical College

Introduction and analysis

THIS survey is based on the circulation of 500 deletion-type questionnaires among 2nd and 3rd year undergraduates at Leicester University. A return of 74% was collated. (368 completed copies). Response comparisons were made between the categories—Male/Female; 2nd year/3rd year; BA/BSc; Special or Honours/General degree. An even spread of questionnaire return was recorded, maximum divergence of response lay at 6% in the BA/BSc case, with 70% and 76% being the respective figures. Students were invited to make additional comments, 55% did so and these everywhere reinforce the inferences drawn from the formal section of the inquiry.

Training:

Student views on training were partially screened by the following questions . . .

(i) Do you think your tutorial : lecture ratio is—*a.* too large/*b.* about right/*c.* too small?

Census

CATEGORY	% SUPPORT FOR ALTERNATIVES		
	a.	*b.*	*c.*
Ma/Fe	5/3	35/50	60/47
2/3	4/4	39/41	57/55
BA/BSc	7/2	60/24	33/74
Sp/Gen	5/3	45/34	50/63

There is obviously a strong appetite for tutorials among all students and this is least satisfied among BSc's, a fact which may well influence their general outlook.

On this and all other questions the widest disparity of views are found to lie between the Arts and Science fields.

(ii) It is both (1) undesirable, and (2) impracticable, that an undergraduate should pursue self-interested lines of inquiry on aspects of the degree course in which he or she is involved.

Do you agree?—*a.* on points (1) and (2)/*b.* on point (1) only/*c.* on point (2) only/*d.* on neither points (1) or (2).

Census tabling is unnecessary here. All categories were unanimous in their views. 93% of all students advocate the desirability of original study as part of a degree course. 39% qualify this by expressing the impracticability of such work in their particular cases.

Assessment:

Two questions were employed to extract ideas on the method of dispensing final awards . . .

(iii) Which do you think provide the best media for determining a students talent?—*a.* essays/*b.* exams/*c.* neither essays nor exams/*d.* both essays and exams.

Census

CATEGORY	% SUPPORT FOR ALTERNATIVES			
	a.	*b.*	*c.*	*d.*
Ma/Fe	14/13	13/3	13/9	60/75
2/3	11/18	9/11	12/9	68/62
BA/BSc	20/10	7/11	5/15	68/64
Sp/Gen	18/9	9/10	12/10	61/71

The 10% difference in B.A./B.Sc. support of alternatives *a.* and *c.* would probably not have arisen had alternative *a.* been extended to include 'other problem solving works'—thus taking special maths students, and the like, into account.

(*iv*) Do you think academic progress at university is—*a.* stimulated/*b.* unaffected/*c.* restricted—by the prospect of examination?

Census

CATEGORY	% SUPPORT FOR ALTERNATIVES		
	a.	*b.*	*c.*
Ma/Fe	58/54	8/9	34/37
2/3	58/55	8/9	34/36
BA/BSc	49/62	6/10	45/28
Sp/Gen	56/57	10/7	34/36

There was marked awareness here of a double standard regarding the interpretatio of 'academic progress'. If the student regards the exam goal as all-important then stimulation is felt. If any kind of original study has primary appeal then a sense of restriction occurs.

It is interesting to view the BA/BSc rift more closely on the interlocking issues of questions (iii) and (iv)—see figs 1 and 2 which show the widest rift as it exists between the categories Ma BA Sp/Ma BSc Sp:

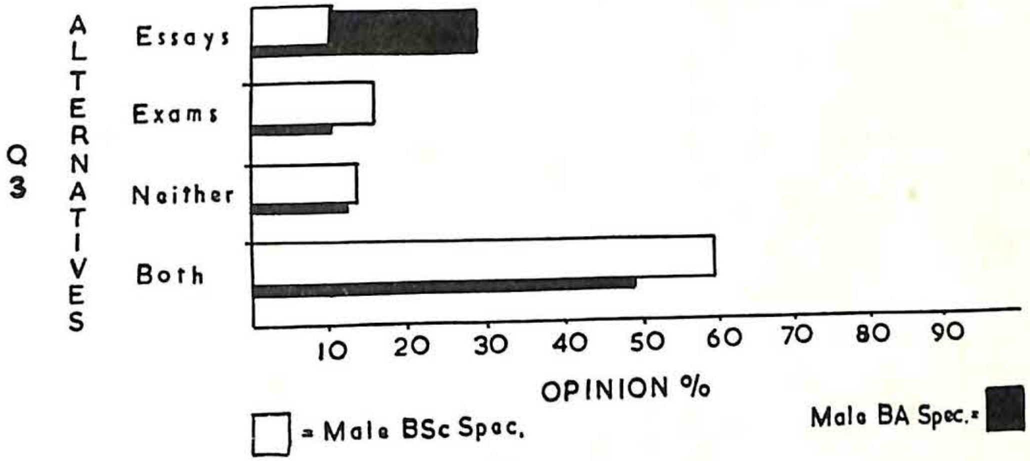


Fig. 1

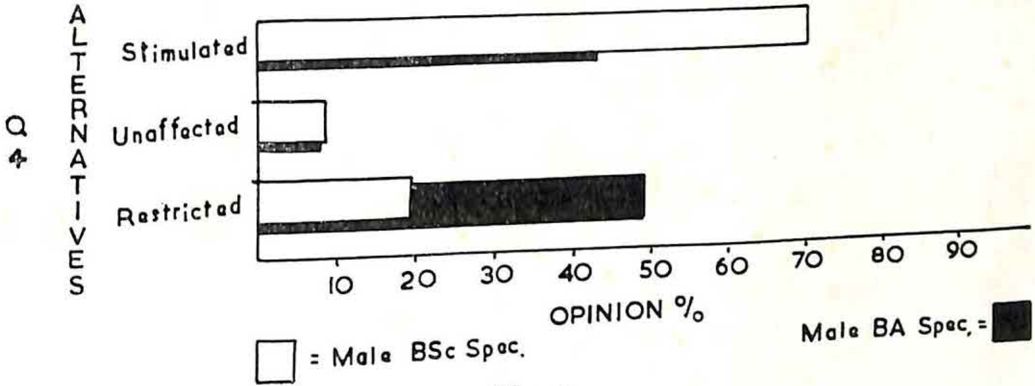


Fig. 2

Although the nature of course subject matter must influence BA/BSc ideas it is rather hard to appreciate that this alone would result in the consistent 'two-culture' pattern so far demonstrated; for under both heads qualitative and quantitative studies are variably mixed. Again, all students show roughly the same progression of exam success in order to achieve undergraduate status, i.e.: 11 + → GCE 'O' → GCE 'A'. Hence it may be, that if the BSc demand for the tutorials which are often totally lacking were met, then such a contrast of views might not be so marked.

Career:

(*v*) On graduating do you wish to—*a*. do research/*b*. teach/*c*. neither teach nor do research?

Census

CATEGORY	% SUPPORT FOR ALTERNATIVES		
	a.	*b.*	*c.*
Ma/Fe	37/14	33/53	30/33
2/3	34/22	35/46	31/32
BA/BSc	22/34	43/37	35/29
Sp/Gen	38/17	32/50	30/33

As might be expected preferences for post-graduate appointments emerged in accord with the alternative courses of study. For some observers the high interest in teaching might be both surprising and encouraging.

Summary and conclusions

A brief review of student ideas and preferences is given on undergraduate conditions of work, on the basis of graduation, and on post-graduate intentions. It may be that an extensive and more detailed survey will show the tutorial system to have considerable impact in modifying student approach to work and assessment. Tutorials and the handling of original topics have high appeal. Whilst exams figure prominently in the minds of most students many alternatives are visualized as separate or combined bases for assessment. Teaching is the most popular of all careers for graduates.

Finally, if it is agreed that all degree training, irrespective of course undertaken, should include a high proportion of tutorial time, a period of original study, and a broad-based 'phased' assessment programme, then many students involved in this survey seem not to receive an adequate percentage of these conditions.

Acknowledgement

During the 1960/61 session the University Education Department gave me every facility for conducting the survey and it is here appropriate for me to insert an expression of gratitude for the help which I received.